Philosophers' Imprint

volume 18, no. 20
september 2018

The Discretionary Normativity of Requests

James H. P. Lewis

University of Sheffield

Why should we ever accede to requests? Is there some common normative basis for the reasons that we act on when we do accede to requests, and if so, what is it? One motivation for trying to answer these questions is the desire to vindicate requesting as a *rational* form of interpersonal reason-giving. This is a necessary piece of the broader vindication of our deeds as those of rational agents. If there is no good answer that explains the normative foundations of requesting, then there is a sort of nihilism hovering in wait. That is, if there are no sound normative grounds for acceding to requests, then perhaps we only ever do so out of socially inculcated habits which themselves cannot be justified. Interestingly, scholars in the field of cross-cultural pragmatics have found evidence to support the view that our tendencies to use requesting as a form of reason-giving are culturally relative phenomena.[1] Perhaps this view itself supports the aforementioned nihilism about requesting: we only make and accede to requests out of culturally contingent habits, not on the basis of normatively grounded reasons. So there is philosophical worth in seeing what defence can be made against that nihilist perspective.

A second, further motivation for thinking about the normative basis of requests comes from its possibly fruitful connection to what some philosophers call the second-person relation.[2] If, in certain circumstances, we can create, at will, genuine reasons for others to act, then what does this say about the kind of relations that we always stand in to other people? One possible thing it might say is that just by virtue of being a person, one has the status in the eyes of others as a source of practical reasons. More than that: as one whose intentions for other people to act can themselves count as good reasons for those others to so act. If it is indeed *just* by virtue of being a person that one has the power to make a request of another, then it might look as though the normativity of requesting is a structural feature of the second-person

1. See for instance (Blum-Kulka, House, & Kasper, 1989), (Economidou-Kogetsidis, 2010), (Huangfu, 2012).
2. I am thinking here in particular of Stephen Darwall (2006), though the interest in the second-person is now much broader, as illustrated by two recent special journal issues (Conant & Rödl, 2014), (Eilan, 2014).

relation. That, I take it, would be an interesting conclusion, and would also have interesting ramifications for moral philosophy. It is not exactly the conclusion that I will defend here, though the intrigue of this hypothesis is part of my motivation.

The thesis that I *am* going to defend is this: a request requires for its efficacy that the person addressed by the request (the addressee) places discretionary value in the person making the request (the addressor). My primary task is to present and stand up for that thought, so I will not have space, other than between the lines, to make conjectures about the consequences of this view for thinking about the second-person relation. This thesis is driven by a need to cover the following three *explananda*, which I believe set the bar for any attempted account of the normative quality of requests. (i) Requests can in principle be the source of *legitimate* reasons; (ii) requests create *new* reasons; and (iii) the reasons that they create are in some sense *discretionary*. To be sure, this triad is, at least at first glance, hard to reconcile.

However, that is the task at hand and it will be undertaken in the following manner. First I will discuss the definition of requesting, making clear how it is distinct from other kinds of interpersonal reason-giving. In the second section, I will set out the notion of a discretionary value, explaining the particular sense in which one could place discretionary value in a person. Third, I will lay out a theory by David Enoch (2011) of what he calls 'robust reason-giving', which is a class of reason-giving that includes requesting. I am sympathetic to Enoch's picture and I will try to illustrate its advantages, but I will also argue that it fails to account for the whole triad of *explananda* when it comes to requests. As such, in the fourth section, I will propose a solution — an adaptation of Enoch's theory to specifically explain the normative structure of requests. There is an apparently compelling objection to my proposal that I will address in the fifth section. If my thesis holds true then it provides a rebuttal of the nihilist suggestion that requests can never create well-founded reasons, and a vindication of requesting as a feature of interpersonal life.

1. Distinguishing requests

I shall define requests as follows:

> **Definition:** A request is an attempt by an addressor to create and communicate a non-obligatory reason for the addressee(s) to perform an action.

A successful request is thus one that succeeds in this attempt. Since my goal in this paper is to consider the reasons that we are presented with in requests and to identify their normative basis, defining what exactly requests *are* is a separate matter, though a crucial one. As such, for the purpose of my main argument, this definition is stipulative. Having said that, I do hope that the concept of requesting under inspection here rings true as a familiar device in the normative play of interpersonal relations. There are, perhaps, a few points in my definition that could be contested, so in this section I will briefly defend two of the most salient of those points. That is, I will defend the ideas that requests should be conceived as *creating* reasons, and that those reasons are *non-obligatory*.

One sceptical perspective from which one might criticise this definition is that of ordinary language. One might reasonably object that when we talk about requests, we are often talking about utterances that do not create reasons but merely state reasons that were already in play; or similarly, it may be that we use the term 'request' to refer to the exchanging of reasons that are obligatory, not non-obligatory. I do not contest that, as we commonly use the term, it does often include these features that are not captured by my definition. I am not providing a definition of the concept as it is used in ordinary language. Rather, the goal here is to define requests as a form of reason-giving with *distinctive normative force*. Specifically, that distinctiveness from other forms of reason-giving derives from thinking of requests as uniquely occupying a quadrant in the chart below (fig.1). Again, whilst this definition is stipulative, I also believe that it captures the heart of the concept of requesting. Thus, when in ordinary language we call

something a request though it does not meet these conditions, I suggest that we are thereby deviating from, and ever so slightly perverting, the true meaning of the term.

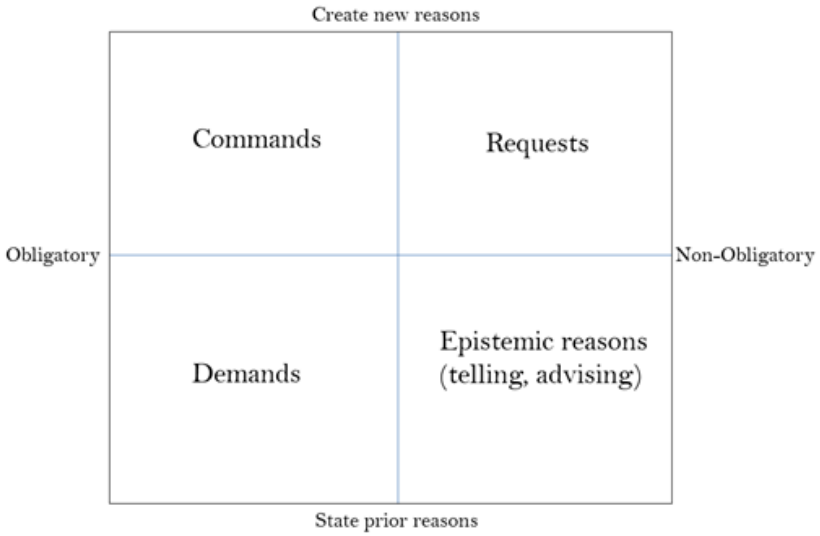


Figure 1

Figure 1 illustrates the conceptual terrain for interpersonal reasons — the kind of reasons that one person is able to give to another person, intentionally, in an act of address, thereby applying some rational force for or against any particular action.[3] The chart carves the terrain along two dimensions: whether the reason presented is thereby created, and whether it is obligatory. As such, the top-left quadrant represents interpersonally given reasons that are both created and obligatory, which I suggest might be the conceptual heart of the notion of a command.[4] Beneath that, in the bottom-left, are reasons that are also obligatory, but which are not newly created. In these instances, the addressor presses their addressee with an obligation that putatively befalls the addressee irrespective of this instance of it being addressed. Again, my suggestion is that this is the conceptual heart of the notion of demand.[5] In the bottom-right are reasons that are also not newly created, but which are not obligatory either. This is just a kind of purely epistemic reason-giving, as when one tells another some non-normative fact ('the bus is coming in 5 minutes') or a normative fact ('the item you have in your hands is very precious'). These are *pro tanto* reasons, considerations which exist, and pertain

3. Note that in categorising ways of giving reasons for action, this chart does not purport to categorise ways of giving reasons for belief.

4. More precisely, I mean two things: that the conceptual heart of 'commanding' is the act of intentionally creating an obligatory reason for another person; and also, broadly, that commanding is a paradigm representative of the class of acts which create obligations.

5. One might be suspicious of this demarcation in Figure 1, on the basis that the border between demands and commands is rather blurrier than I am suggesting: that sometimes demands create new reasons. Let me try to allay this concern through an example. Consider a group of employees in a factory that produces supermarket sandwiches, who, after a breakdown in official negotiations, down their aprons and march to the management corridor of their employer's head office, where a spokesperson for the workers issues a demand to the employers that they introduce a decent workplace pension scheme. One way to interpret this putative demand would be to understand it as stating an obligation that was already in play, and thereby holding the addressee to account, much like the rebuke one might give retrospectively by saying 'you should have provided a pension scheme for your workers'. Indeed, the spokesperson's utterance of the demand could at the same time be an instance of epistemic reason-giving, making the employers aware of this normative fact that they had hitherto overlooked. Alternatively, one could understand this 'demand' as being, in fact, a command in disguise. That is, the spokesperson may be creating an obligation simply by stating it, invoking their authority as representative of the people on whom the employers depend. Either way, it is apparent that the distinction between demands and commands is a clean, not a blurry one: either the obligation that the utterance purports to present was already there, or the utterance purports to create it. Little hangs on the terminological dispute over what we use the words 'demand' and 'command' to refer to. The point for my purposes is that these terms can be used to capture two *discrete* normative moves.

to the addressee, irrespective of this particular interpersonal address. Finally, then, in the top-right quadrant are non-obligatory reasons that are newly created in the act of being addressed to one person by another: requests.[6,7,8]

An example here may be helpful in illustrating the extent to which this stipulative definition is revisionary of the ordinary concept of requesting, but also the extent to which it nonetheless succeeds in capturing the core of that concept. Consider Bronwen, a bus mechanic who has recently taken on young Mair as an apprentice. Bronwen is kindly and warm-hearted towards Mair and I suspect that most people would ordinarily take her to be making a request when she asks Mair, 'Please would you change the oils of the Leyland National that came in yesterday?' But if Mair were to treat this reason that she has just been given as *discretionary* (a notion that will be elaborated upon shortly), then I fear that Mair would have misunderstood the situation quite badly. If she simply chooses not to change the oils of the Leyland National then she will have something to answer for, and Bronwen could legitimately express at least a little irritation towards her.

Similarly, suppose that a few hours later, Bronwen notices that Mair has not gone near the Leyland National, and so she reiterates her earlier utterance. Again, there is an ordinary sense in which the reiterated utterance would commonly be thought of as a request. But woe betide Mair were she to think of utterances like these as creating *new* reasons that were not in play prior to their being uttered. Whilst we may ordinarily refer to both Bronwen's first and second utterances as requests, on my stipulative schema, they are a command and a demand, respectively. The first created a reason for Mair to change the oils that she didn't have prior to having been asked, but a reason that is obligatory, rather than discretionary. The second did not create a reason at all, but merely reminded Mair of an obligation that she had been given, where the reminder itself was an act of holding Mair to account to that obligation.

Whilst being revisionary, though, there is nonetheless some purity to the notion of requesting that is advanced here. The claim that the given stipulative definition captures the conceptual heart of requesting rests on the further intuition that although we may ordinarily refer to Bronwen's speech acts as requests, we may at the same time acknowledge a sense in which they were not *really* requests. I make this sentiment explicit by claiming that speech acts are only *really* requests when they attempt to create non-obligatory reasons, and stipulate that hereafter in the present discussion, the term 'request' will be reserved for the real deal, so to speak. Thus, when appeal is made below to intuitions about requests, these appeals seek to draw on intuitions about only this central subset of what the term 'requests' often includes.

I have thus put aside the ordinary language objection on the basis that what is required here is a stipulative definition that captures the distinctive normative force of requesting. But that will not satisfy a detractor who maintains that not only does the word 'request' commonly refer to reasons in the other three quadrants of the chart here, but moreover, that the quadrant that I am designating to requests is either unnecessary or an impossibility: there can be no such things, and/or we need not think that there are such things, as speech acts that create non-obligatory reasons. In a way, the thesis that I want to defend in the later sections of this paper speaks to this challenge by explaining how

6. One of the few philosophy papers that expressly sets out to address the normativity of requests provocatively argues that requests *do* create obligations. That is Cupit (1994), who thinks that requests appeal to obligations that are grounded in the commitments of the agent. Despite appearances, this view is interestingly similar to the one that I will develop here, though I do ultimately disagree about the characterisation of the resulting reasons as obligations, rather than discretionary reasons. Also noteworthy is that Cupit (1994, p. 449) agrees with the other half of my definition: that requests "generate" new reasons.

7. Defining requests as giving non-obligatory reasons is concordant with Lance and Kukla (2013, p. 460): "The [normative] output of a successful imperative is an obligation on the part of the person ordered to do what the speaker ordered her to do. The output of a successful request is that the target now has a specific sort of reason to do what was requested, but it is essential to the notion of a request that this reason is not an obligation."

8. This definition of requests also conforms to Raz's understanding of requests as a kind of content-independent reason-giving (Raz, 1988, pp. 36–37), with which Owens (2012, p.86) also concurs.

there *could* be a normative basis for such reasons. But a more immediate response to this challenge can also be made.

The challenge in question, to be clear, rejects my stipulative definition of requests on the following basis: that the other three quadrants of the chart between them exhaustively explain the reason-giving that goes on in requesting; that no recourse is needed to the idea of a newly created non-obligatory reason. I think this is not true. The fact that the kinds of reason-giving categorised in those other three boxes are inadequate to explain at least some of what we conceive ourselves to be doing when we make and accede to requests, can be seen through an example.

Consider two friends, Sioned and Ffion. Sioned is mounting an election campaign and she wants Ffion to help as her campaign manager. Committing to the campaign would constitute a substantial sacrifice for Ffion as it will be stressful, and for the course of the campaign it will take a lot of time away from her own work, her family and her other engagements. Suppose that Ffion knows perfectly well that Sioned wants her help: indeed, everybody knows it. But because of the extent of the sacrifice that it would entail, Ffion has not voluntarily offered her help to her friend. For some time, Ffion knows that Sioned desires her help and Sioned knows that Ffion knows this too, but, somehow, she cannot bring herself to ask for help: partly out of pride, partly out of reluctance to burden her friend, partly in the hope that an offer will be forthcoming from Ffion anyway. But it is not, so the time comes and Sioned confronts the awkwardness that has arisen between them with a request: she explicitly asks Ffion whether she would commit to helping Sioned's election bid in the role of her campaign manager. This, I suggest, is a request which presents a non-obligatory reason for action. But moreover, the request itself has altered the normative situation. I suggest that it has done so by creating a reason that was not present before.

All I mean to appeal to here are two intuitions about this case (and therefore about others like it). The first intuition is that the reason presented by the request is not an obligatory one. Of course friendships are relationships that can sometimes generate obligations, including obligations to help one another.[9] But being someone's friend does not mean *always* being obliged to help them, regardless of what the help is needed for, or of the cost that helping would incur. Sioned's request is supposed to be an example of one of those deeds which Ffion is not obliged to do: Sioned is not at risk of being harmed and the costs of helping are substantial. As such, the kind of reason-giving that is going on in this instance cannot be anything on the left-hand side of the chart in figure 1, since the reason is non-obligatory.

On the other hand, the second intuition is that the making of the request is a normatively significant moment. Whatever the act of uttering the request does, it has some kind of impact on the balance of reasons for and against helping. That is to say, from Ffion's perspective it makes a difference that Sioned asks for her help. Since in the example it is stipulated that prior to the request being made, Ffion already knows full well of Sioned's desire for her help, it seems that the normative difference that the speech act of the request makes cannot be an epistemic matter. It cannot be the case that Sioned alters the normative situation through some epistemic reason-giving, telling Ffion about her desire for the latter's help, because Ffion already knows all about it.[10] As such, the normative role that the request is playing must be more than merely epistemic. Somehow or other, this utterance is creating a new reason for Ffion to commit to the campaign itself.[11]

9. For a discussion of how friendships and other particular relationships can generate special obligations, see (Jeske, 2008).

10. As I will discuss in section 5 below, some theorists — particularly in the domain of speech act theory — do think of requests as functioning by expressing desires. Searle (1969, p. 66) is the precedent for this.

11. One objection here to the claim that requests create new reasons, would be to point to cases where the addressee seemed to have more than sufficient reason to perform the action in question prior to being asked, where the addressor could legitimately say 'I should not really have to ask!' However, in such cases as this, it is at least *possible* that the request does indeed add a new reason to the existing pile. It is exactly the fact that the existing pile *without* a further reason had not already moved the addressee to action that has drawn the addressor's ire.

When added to those of which she was already aware, this utterance could conceivably be sufficient to tip the balance of reasons for Ffion in favour of making the commitment to her friend. This prompts the question that I will try to answer later on, of *how*, without any appeal to authority, it is possible for people to simply create reasons of this sort for others.

Despite having said earlier that the definition of requesting in this paper is not intended as a descriptive attempt to capture the ordinary language meaning of the term, I have nonetheless addressed the relation of my stipulative definition to the ordinary-language conception. Whilst in everyday speech we may commonly use the term to refer to other forms of reason-giving, I have argued — through the example of Sioned and Ffion — that the everyday notion must at least *include* the kind of reason-giving that is under inspection here. In the next section, I will elaborate on the idea of having reasons that are discretionary, and how other people could be the source of such discretionary reasons. This will lay the ground for my actual account of the normative basis of requesting, which I will present later.

2. The discretionary value of persons

What it is to value something is a troublesome question. But whatever else it is, valuing X is having a favourable attitude towards X such that X can be the source of reasons for certain actions. Thus, if you value this photograph of your grandmother, as well having an array of beliefs about, and emotional attitudes towards it, you will also be disposed to act in certain ways regarding the photograph. You may, for instance, be prone to lurch to catch it if you were to see it falling; or perhaps you would be inclined to act to ensure that it retains a prominent, visible position on the mantelpiece — whatever.[12] In this section I want to set out the sense in which valuing another person can make them the source of reasons in this sense (that is, in the internalist sense that they would not be such a source of reasons were it not for one's attitudes towards them). It will also be helpful here to clarify what bearing this kind of internalism about this class of reasons has on other important matters, namely, the (non-)obligatory quality of the reasons, and their relation to external reasons and to moral reasons.

A first important thing to note is that we are morally obliged to value others, to a certain extent. Typically, we might express this by saying that we are obliged to respect others, or some similar thought. It seems plausible to me (though nothing hangs on this here), that our obligation to respect all people does not simply pertain to performing the deeds of respecting their dignity. Rather, the obligation also pertains to the attitudes that we hold. We must actually *respect* others' dignity, which means, *inter alia*, believing that their dignity is worthy of respect and perhaps harbouring at least some minimal degree of emotional connection with their dignity — a disposition to regard threats to their dignity with concern, for instance. Holding the attitude of respect towards others is holding an obligatory interpersonal valuing-attitude.[13]

Indeed, respect might be an attitude that we are each obliged to hold towards everyone else, but there are other obligatory interpersonal valuing-attitudes, ones which are specific to one's particular relationships with others. Often, for instance, one is obliged to hold the attitudes of deference towards one's elders, of sympathy towards one's young children, of solidarity with one's friends and comrades. In entering into relationships with others, one accrues obligations to hold certain attitudes towards them. And the obligatoriness of such attitudes is not undermined by the fact that forming those relationships with those people was not, in the first place, something one was obliged to do.

But not all interpersonal valuing-attitudes are obligatory. Consider the affection you might have for your nieces, the admiration you

12. See (Scheffler, 2010, pp. 27–28).

13. I do not think that one is obliged to *always* hold this attitude towards *everyone* in the world. Whilst there may well be obligations that we all owe to absolutely everyone, the obligations to hold certain attitudes towards people only emerge when you are (or should be) aware of those people.

might have for a colleague, pity for an unfortunate stranger, endearment to a charming one, the lofty esteem that you might have for an able rival. These attitudes are all discretionary. No one is entitled to demand them, or to blame others for preferring not to hold them. In fact, for any valuing-attitude that can be legitimately expected, one can imagine the possibility of holding that attitude more intensely than is expected. The fact that one is obliged to hold a valuing-attitude to a certain extent implies that whether to hold it to a greater extent is a matter of discretion.

Significantly, there is a difference between valuing a quality, or set of qualities that someone has, on the one hand, and on the other, valuing *them*.[14] When discussing interpersonal valuing in requests — which I will get on to in the next sections below — the kind of valuing must necessarily be the latter kind, not the former. Various qualities may play a role in leading us to place value in others, or in vindicating in our own eyes the value that we already do place in others. But our interpersonal relationships are characterised by the way in which we value persons themselves, not merely certain aspects of them.

A consequence of the fact that it is people as individuals that are the objects of our interpersonal valuing-attitudes is that the reasons that arise from those attitudes are reasons to act *for others*. Thus, when my admiration of you manifests in giving me a reason for some kind of action — to help you in your endeavours, perhaps — I thereby have a reason to act *for you*. Doing someone a favour — which is a way of acting on an interpersonal valuing-attitude — entails doing something *for* that person herself. What this kind of inter*personal* valuing attitude contrasts with would be a kind of valuing attitude that pertains only to particular traits and qualities: fondness for Harriet's dry wit but not fondness for Harriet *per se*.[15] The point here is one that will be relevant later: that there is a species of discretionary valuing-attitudes that play a central role in our social lives, which have as their objects people, as such, rather than merely the valuable qualities that people sometimes bear.

At this point, it is worth noting the distinction between being obliged to do something, on the one hand, and having *reason* to do something all things considered.[16] That is, it might be said that one 'rationally ought' to do that which one has most reason to do, all things considered. But in distinguishing a set of attitudes that one must hold, from another which one may hold at one's discretion, I am not making any claim about what one rationally ought to do, or where rationality allows for some discretion.[17] Rather, the sense of obligation at play here is to do with what it is morally right and wrong for us to do.[18]

Moreover, the obligatoriness of a valuing-attitude finds expression in the obligatoriness of the actions that express that valuing-attitude. I have already mentioned the strong connection between valuing something and treating it as a source of practical reasons. This connection

14. The difference I want to capture matches Darwall's (1977) distinction between appraisal respect, and recognition respect, respectively. Having said that, I do not mean to endorse what Darwall thinks are the bounds of the recognition respect that persons can enjoy from one another, where "there can be no degrees of recognition respect for persons" (ibid., p.46). Unlike Darwall in that paper, I am concerned here with *non*-moral interpersonal valuing, within which sphere it is possible to value some people more than others, in a way that is not at all a matter of appraising their virtues.

15. The valuing I have in mind could be characterised as valuing someone *de re* and not *de dicto*. For a discussion of related matters see (Kraut, 1986, esp. p.423).

16. Thanks to Bob Stern for pressing this distinction.

17. As it happens, I do also in fact think that there is such a thing as rational discretion *and* that interpersonal valuing-attitudes that are discretionary in the deontic sense are also discretionary in the sense of there being no determinate all-things-considered set of attitudes that any given person rationally-must adopt. But for present purposes, the notion of rational obligation is not relevant.

18. What characterises obligations is a matter of some controversy. One influential account is Raz's view of obligations — or mandatory reasons — as involving a second-order 'exclusionary' component that instructs the disqualification of competing first-order considerations (Raz, 1999, pp. 73–76). Other theories of obligation define the concept in terms of the kind of accountability that it implies (Darwall, 2006, chpt.5). For my purposes, I need not endorse one account or another, so long as they are all compatible with a general thought that obligations "always give agents conclusive reasons for acting that outweigh or take priority over any potentially competing considerations" (Darwall, 2006, p. 26).

illustrates the divide between obligatory and discretionary interpersonal valuing-attitudes. Suppose that Charlene is obliged to hold an attitude of deference towards her professor, Dominique. If she fails to hold that attitude she will be doing something wrong. This failure may manifest itself in certain deeds — an insufficiently deep bow, a lacklustre display of courtesy, perhaps — and by extension, these deeds too are wrong. By contrast, there is nothing wrong about failing to feel heartfelt affection for someone, or genuine admiration, or real pity. It is intrinsic to the very notions of these attitudes that one cannot be obliged to feel them. The true sentiment of affection can only be an organic sentiment, one that arises naturally and not out of duty.[19] The same goes for the 'true sentiments' of other interpersonal valuing-attitudes.

The notion of a valuing-attitude being discretionary is important. An attitude is discretionary just when it is not obligatory. And on the picture of practical reasoning that I am assuming in this paper, an attitude is non-obligatory just when no one is entitled to react with anger to one's holding or failing to hold the attitude. When we are unconstrained by duties, in this sense, we must exercise discretion over our conduct: we take ownership over which values to invest ourselves in, and over how we weigh those values against one another. This thought too will be relevant later in the discussion of requests. Since requests do not create obligations, I will claim that they must make an appeal to their addressees as agents who have this kind of discretion.

In this section, I have been trying to express an idea that I think comes naturally when we think about the reasons that we have to act in the interests of others. That is, I have tried to establish — in line with common intuitions — that there are such things as discretionary reasons to act for another person that stem from discretionary attitudes of valuing that person. Before moving on, there are two noteworthy features of the general picture of practical reasoning to emphasise. The first is that, as mentioned above, the reasons that I have been chiefly concerned with are 'internal' reasons. That is to say, they are normative reasons that make essential reference to some aspect of the motivational set of the agent for whom they are reasons.[20] However, that does not mean that the theory that I am advancing represents a partisan position on the debate in meta-ethics between internalists and externalists about moral reasons; it doesn't. It is entirely compatible with the view that people have some reasons that depend on their own values, to also think that they may also have some other reasons — moral obligations, perhaps — that are external to their own set of values. I thus remain neutral on that question. Having said that — and this is the second noteworthy point — the kinds of discretionary other-regarding reasons that I have identified *could* have some moral significance. Specifically, when moral philosophers talk of something being the wrong or the right *kind of reason* for someone to do something, the discretionary interpersonal reasons discussed here might seem relevant. It seems plausible, *prima facie*, that acting out of a genuine, discretionary heartfelt desire is very much the *right* kind of reason to act, even if the action itself is something that one is morally obliged to perform. So this is just to note that whilst the picture that I am advancing is neutral between competing moral and meta-ethical theories, it may have some interesting consequences.[21]

3. Enoch's account

It is now possible to return to the goal of explaining what, if any, normative force there might be to the new, non-obligatory reasons that are presented in requests. David Enoch (2011) has devised a sophisticated account addressing this issue. Here I will offer a sketch of how Enoch proposes to explain the normative power that people have to give practical reasons 'robustly'. There is, though, a problem with this explanation when it is brought to bear on requests. I will try to

19. For a further defence of the view that there cannot be a duty to love, see (Driver, 2014). Not everyone holds this view, however, see (Liao, 2006).

20. For an elaboration on this kind of understanding of internalism, see (Markovits, 2014); for its classic source, see (Williams, 1981, p. 102).

21. Interesting though such consequences may be, discussing them properly must remain a matter for another time.

illuminate the difficulty that Enoch's view has in accounting for the discretionary quality of the normativity of requests. In the following section below (section 4), I will propose an amendment to the view that enables it to overcome this problem.

The primary concern driving Enoch's inquiry is the explanation of a phenomenon broader than merely requesting. The motivating question is rather, "if, as seems likely, 'reason must constrain and guide the will', how is it that we can create reasons at will" (Enoch, 2011, p. 1)?' The sphere of intentionally created reasons includes commands and promises, and to the entire domain he gives the label 'robust reason-giving'. But Enoch thinks that requests are the paradigm form of robust reason-giving by virtue of their simplicity relative to these other apparently more complex phenomena, where authority is involved. As will become apparent below, requests have complexities of their own. But the theory of robust reason-giving in its general form is still of use. According to that view, all practical reasons that one person can give to another can be categorised into two sets. On the one hand, they could be a kind of merely epistemic reason-giving (the sort of thing found in the bottom half of the chart in Figure 1 above). Alternatively, if they are doing something more than merely telling or advising the addressee about some prior existing reasons, then, Enoch thinks, they must be *triggering* a reason. As such, since these reasons by stipulation are more than merely epistemic reasons, they must be instances of triggering-reasons.

It may be helpful to elaborate on this point. A robustly-given reason — such as a command or a request — does its normative work not simply by trying to reveal to the addressee what reasons there are for them to act, but by in some way *changing* what such reasons are. But practical reasons are not the sorts of things that can be merely willed into existence wantonly. You cannot, for instance, make it the case that a stranger should arduously undertake to do your bidding, merely by deciding that they should. Rather, Enoch (2011, p. 9) infers, these reasons work by *realising the non-normative antecedents* of conditional reasons that hold true independently.

One example of this can be seen in the case of commands, which are a species of robust reason-giving. When the sergeant commands one of her officers to quick march to the barracks, she triggers a conditional reason, by realising its antecedent. That conditional reason must have the form: 'If commanded to do so by the sergeant, then the officer has (obligatory) reason to quick march to the barracks.' And the same story applies to requests. When Sioned requested Ffion to help her with the campaign, she triggered something like the following conditional reason: 'If requested to do so by Sioned, Ffion has (non-obligatory) reason to help with the campaign.' Such reasons can be made when the relevant conditional reasons are true; conversely, successful robust reason-giving implies the truth of the prior conditionals (Enoch, 2011, p. 10).[22]

This is not the whole of the account, however. So far, no space has been made for the difference between robust reason-giving, on the one hand, and, on the other, the variety of other ways in which non-normative circumstances can be manipulated so as to trigger conditional reasons. Enoch (2011, p. 4) gives the example of the neighbourhood grocer raising the price of milk. By doing so, we can suppose that she triggers (again, by realising the antecedent of) a prior conditional reason, one of the following sort: 'If the price of milk at this shop is above X amount, it is too expensive, so you should not buy milk here.' Here, the mechanism by which the grocer *inadvertently* gives you a reason not to buy her milk looks identical to the mechanism by which the sergeant creates a reason with her command, or that by which Sioned creates her request. This is inadequate since it certainly seems that the normative power of reasons like those presented in requests simply

22. Following a line of argument by Mark Schroeder (2014), Enoch (2011, p. 11) acknowledges the conceptual possibility of robust reason-giving that does not trigger a prior reason but really creates a wholly new reason. Thus, divine command theorists may believe that the obligatoriness of a command *consists* in God having commanded it. On such a view, God gives reasons that do not rely for their force on the truth of prior conditional reasons. But in the present paper I am concerned only with reasons exchanged between ordinary mortals, which, contrarily, *must* be grounded in prior, conditional reasons.

have more to them, so to speak, than such incidental reason-giving as the grocer's price change creates.

To address this, the account of robust reason-giving must incorporate the role played by the *intentions* of the parties to these exchanges. In requests and commands, the reason that one attempts to make with one's utterance depends on the addressee recognising one's intention for this utterance to give them a reason. Enoch's exact formulation of this thought summarises the account (Enoch, 2011, p. 15):

> One person A attempts to robustly give another person B a reason to Φ just in case (and because):
>
> (i) A intends to give B reason to Φ, and A communicates this intention to B;
>
> (ii) A intends B to recognize this intention;
>
> (iii) A intends B's given reason to Φ to depend in an appropriate way on B's recognition of A's communicated intention to give B a reason to Φ.

There are several considerations to be discussed in relation to this proposal that I shall leave aside here. For the purposes of my argument, I shall assume that the formulation of robust reason-giving set out here sufficiently explains the general manner in which intentions are relevant to the class of normative interpersonal interactions that are at issue — including commands and requests. In what follows, I will focus in on requests and inspect the status and normative quality of the conditional reasons — those that are required for a request to be successful, whose antecedents are made-true by the uttering of a request.

At this juncture, I would like to raise a problem, or rather, to raise again the problem with which I began. How does the theory of robust reason-giving fare at dealing with the three *explananda* of requesting? The first *explanandum* was that requests can in principle be the source of *legitimate* reasons. One worry in this regard might be that in acceding to requests we act out of socially inculcated habits, rather than on the basis of well-grounded reasons. Another concern is that requesting must always be some kind of coercion, since legitimate reasons cannot simply be willed into existence, out of thin air. The theory of robust reason-giving addresses these concerns. It does so by revealing the role of prior conditional reasons that are brought into play in requests. Those are reasons of the form, 'Person A should Φ, if requested to do so by person B.' These conditional reasons are not willed into existence; they are in some sense there already, before — or at least *at* — the moment of a request being made. By positing the existence of such conditional reasons, the theory can explain in principle how acceding to a request could be justified by reference to a legitimate reason. Acceding therefore need not be thought to be a response to a social convention or being bent coercively by the mere will of the requester.

The second *explanandum* was that requests create *new* reasons. This is the feature towards which the theory of robust reason-giving is primarily addressed. Enoch is motivated by a suspicion of the mysterious-sounding notion of reasons — which bind our wills — coming into existence at the mere whims of agents. But this mysteriousness is played off against the phenomenology of requests. As the example of Sioned's request illustrated earlier, it seems certain that it is possible for requests to make an impact on the normative terrain — to *do* something, that is — even when all the relevant normative and non-normative facts are known, so they cannot be *doing* anything epistemic. Robust reason-giving explains this *doing* as kind of a triggering. By making-true the antecedent of a prior, conditional reason, requests manipulate the non-normative circumstances in such a way that the addressee has a reason that they didn't have before. (That reason is just this: that a request has been made of them.)

The third of the triad of *explananda* was that the reasons that requests create are in some sense *discretionary*. How does the theory of robust reason-giving account for this discretionary quality? This is

where I think the problem arises: I do not think it can. For a request can be made only if there is a prior conditional reason available for it to trigger. In other words, I can only request you to Φ if it is the case that 'you have reason to Φ, if I request you to do so'. If such a prior conditional is not true, then the request will fail. It will fail not just to persuade the addressee conclusively to Φ, but even to alter the balance of reasons at all. But on the other hand, if the prior, conditional reason is true, and it is triggered, then the addressee simply has a reason to Φ, and it is not clear where the discretionary quality enters in.

This is quite a serious concern. It is a fundamental quality of requests that acceding to them is distinctively a matter of discretion. Asking someone to do something is an interesting, special form of reason-giving precisely because in so asking, one intends for the other to treat this request as a reason, but not for them to treat it as itself conclusively instructing them. We ask them to Φ, and thereby acknowledge that whilst our wishes are clear, the matter of whether to Φ or not is up to them. Despite the fact that Enoch thinks of requests as the paradigm of robust reason-giving, that theory lacks any conceptual resources to account for this defining discretionary quality.

It might be thought that Enoch's account as it stands can accommodate the discretionary quality of requests simply by distinguishing them from commands.[23] That is, the reasons presented by requests, rather than being obligatory, are merely *pro tanto*: they are 'first-order' considerations that favour certain actions, but they could just as well be outweighed by more pressing reasons that speak against those actions. Obligations should not be outweighed in this way. Obligations purport to provide conclusive reasons. Thus, the discretionary quality of requests might be thought to consist simply in the fact that they do not purport to provide conclusive reasons.

But this line of defence, though initially tempting, fails on two fronts. First, it simply begs the question. It is true that to be merely *pro tanto*, rather than obligatory, is part of what it is for a request to be discretionary. But merely claiming that requests trigger conditional reasons and those reasons are merely *pro tanto* is not enough. The question is *how* can one person issue another with a merely *pro tanto* reason to do as they ask? In the case of commands, the prior, conditional reasons ('if the sergeant commands, then the officer should Φ') are made-true by the authority that the addressor has and the reasons are grounded in the normative grounding of that authority. These normative powers are often taken for granted: the parents' authority over their teenage child plausibly stems from their parental responsibility; the employer's authority over her employee stems from an explicit contract; the restaurant customer's from an implicit one; and so on. To be sure, in any given case, it may be far from straightforward to determine whether the putative authority is in fact well grounded. What is straightforward, though, is that *when* reasons are robustly given in commands, they are *always* grounded in this kind of authority. In requests, though, there is no authority at play. So what, on Enoch's picture, could ever make it true that person A has a merely *pro tanto* reason to Φ if requested to do so by person B? This is the question that the theory is so far ill-equipped to answer.

The second front on which that tempting line of defence fails is its characterisation of the discretionary quality of requests. One of the distinctive things about requests is that, at some level, they appeal to addressees to *choose* what to do: to make a choice between competing values, rather than simply calculating what they have most reason to do. Or, put another way, when we do appeal to others to make a certain choice between competing values, it is possible for us to make such appeals by requesting. But insofar as requesting is understood merely as the simple triggering of a reason, this aspect of the phenomenon remains mysterious. The general theory that Enoch has articulated has no resources to explain how requests can appeal to their addressees to choose between such competing paths where only they, the addressees, have the authority — the discretion — to make that choice.

23. Thanks to Daniel Viehoff for pressing this point.

4. A proposal

In response to these difficulties, I have a sympathetic proposal to amend the account of robust reason-giving as it applies to requesting. My suggestion is that the prior reasons that are triggered by requests must always have not one but two conditions. Besides being conditional on a request being made, they must also depend on the addressee placing some discretionary value in the addressor. Accordingly, those prior, conditional reasons take the following general form:

> Person A has a reason to Φ *if* [condition (i)] requested to do so by Person B *and if* [condition (ii)] Person A places sufficient discretionary value in Person B.

On this view, then, the normativity of requests is keyed to interpersonal valuing-attitudes. Specifically, it is keyed to a set of valuing-attitudes which are a matter of an agent's discretion: these are attitudes that one is not obliged to hold. Specifying this point, therefore, introduces sufficient conceptual resources to explain the discretionary quality of the reasons presented in requests.[24]

My proposal bakes in the discretionary character of the reason at the level of the prior, conditional reason. A consequence of doing so is that if that discretionary valuing-attitude is not held by the addressee — if condition (ii) is not met — then the request fails entirely and does not create a reason at all. And this outcome is one that might seem problematic. That is, one might think that even when a request is made by a contemptuous fiend — a person towards whom one holds no discretionary valuing-attitude whatsoever — a reason might nonetheless be *created*.[25] The thought is that the reason might be created just as the fiend intends even though in the addressee's deliberations it has insignificant normative weight, or is dramatically outweighed by countervailing considerations against acting *for* the contemptuous addressor.

But this problem does not arise if one keeps in mind the structure of requesting as a distinctive normative operation — a structure that depends on appropriate mutual acknowledgement of the intention to create a reason. Recall that in the theory of robust reason-giving, to which my proposal is an amendment, requests create reasons only when the addressor intends the addressee's given reason to Φ to depend in an appropriate way on the addressee's recognition of the addressor's communicated intention (to give the addressee a reason to Φ). On my proposal, the addressor intends to trigger a reason which itself depends on the addressee's discretionary value-outlook. As such, the addressor intends the request as an appeal to an item in the addressee's own discretionary value-outlook. If there is no such item, if the addressee does not place any discretionary value in the person of the addressor, then the request fails to create a reason. Moreover, it fails to create a reason *even by the addressor's own lights*. Of course, the fiend may succeed in coercing the addressee, or the fiend may have the authority to command her, or it is even possible that the contemptuous fiend can reveal his desire (epistemically) for the addressee to do his bidding, and that mere desire may give the addressee a reason. The fiend may be successful in creating reasons in all these sorts of ways. But without the addressee holding discretionary value in their addressor, the latter cannot create a reason in the normatively distinct sense of requesting.

24. A question that might be raised to my view (and indeed has been, by Alfred Archer, to whom I am duly grateful), is whether the reasons that are thus created by requests are free-standing considerations that favour Φ-ing, or whether they can only ever play an accompanying role to other reasons that must also be at play. Specifically, does the reason that is created by the request (qua request) depend on there being a favouring reason that stems simply from the existence of the addressor's desire for the addressee to Φ? In the terms of Jonathan Dancy's work on the different sorts of practical reasons that there are, this is the question of whether requests create 'favourers' or 'intensifiers' (Dancy, 2004, pp. 38–43). My answer is that the reasons created by requests are stand-alone favourers. When we hold an interpersonal valuing-attitude in another we endow another with the power to create reasons by requesting. Conceptually speaking, I do not see why it should be impossible to endow someone with this power, whilst for whatever reason not treating their very desires as themselves the sources of practical reasons.

25. I am very grateful to Glenda Satne for pushing me on this point.

As a further illustration of the theoretical worth of the proposal being made here, consider the case of entreaties. I shall use the term 'entreaties' to refer to a subset of requests in which, prior to the request being made, condition (ii) is not met: the addressee does not yet place sufficient discretionary value in their addressor to grant them the standing to make the request.[26] In entreating, one attempts to trigger *both* conditionals of the prior reason. That is, the addressor appeals to their addressee to actively place discretionary value in them — in the person of the addressor — and simultaneously to request, on the basis of that discretionary valuing-attitude, that the addressee undertakes some action.

Suppose that Carrie and Anita are strangers to one another. Carrie is walking down the street on which Anita lives and she urgently wants somewhere to hide, but to tell anyone why she needs to hide would risk endangering her confidant. She knocks on the door of a house on the street and Anita answers. Carrie asks whether she can come in, without offering any explanation. We might imagine that she asks whilst looking directly into Anita's eyes. In the moment before this exchange, if Anita had been asked 'would you consider the request of a stranger to come into your house, without explanation', she would have said no. Anita would not have granted a stranger even the standing to make that request — not without some explanation. But in the moment of the entreaty, Carrie implicitly appeals to Anita to take up some kind of valuing-attitude towards her. This could be admiration, affection, pity, some kind of endearment. (As I mentioned in section 2 above, these particular feelings are only points in an indefinite range of favourable interpersonal attitudes that one can hold, or not, at one's discretion.[27]) It is a presupposition of Carrie's entreaty that the following conditional reason is true: Anita has reason to let Carrie into her house, if Carrie asks, and if Anita places sufficient discretionary value in Carrie.

A theoretical advantage of my proposal is that it equips the theory of robust reason-giving with the capacity to explain the normativity of exchanges like Carrie and Anita's, of entreaties in general. This advantage is pertinent because, as it seems to me, the boundary between entreaties and ordinary, run-of-the-mill requests is fluid and often difficult to identify. A run-of-the-mill request, let us say, is one in which condition (ii) is met already, prior to the request being made. All it takes is for the request to be uttered and a reason will have been created for the addressee, without any alteration in anybody's value-outlook also being required. For example, Gwen and her younger sister Cat are being looked after by their babysitter, Wynn. Gwen is extremely enamoured of Wynn — thinks the world of her — so when Wynn asks Gwen to go and read Cat a bedtime story, there really is no question of whether she places sufficient discretionary value in Wynn for the request to be reason-giving. This is a clear-cut case of a run-of-the-mill request. But I suggest that often, depending on the demandingness of the action that is being asked for, things are less clear. When we make requests, it seems that we often implicitly appeal to our addressees to value us — we seek to convey our worthiness of pity or esteem, or whatever, as a way of bolstering the reason that our request attempts to provide. On my view, these appeals may sometimes be requirements for the request to succeed in creating a reason at all. If the border between entreaties and run-of-the-mill requests is indeed as hazy as this, then any account of requests should be capable of explaining *at least* how there could be such a hazy border. Enoch's account cannot, since the interpersonal valuing attitudes that are appealed to in entreaties play no role in his theory of requests. But my proposal explains the possibility of the hazy border, as well as giving an account of the normativity of the terrain on both sides of that border: of both entreaties and run-of-the-mill requests.

26. In giving this specific meaning to the term entreaties, I am following Lance and Kukla (2013, p. 474): "[A]n entreaty is a meta-call: it calls someone to grant the caller an entitlement to make certain kinds of claims that the caller is not yet in a position to make."

27. Having said that, Cristina Roadevin has pointed out to me that some requests make their appeals to *particular* interpersonal valuing attitudes. For instance, the lover's request is appropriately granted out of love, affection, but not out of, say, pity.

In this section I have presented an amendment to the theory of robust reason-giving. The amendment helps to explain the sense in which the reasons presented in requests are discretionary reasons. I will now raise an objection to my proposal in the form of a competing explanation of the discretionary quality of requests, one that is popular in the way that speech act theories think about the matter. I will argue that this competing explanation fails to address the normative questions that my view sets out to confront.

5. The pragmatics of discretion[28]

A central merit of the account that I am proposing is its capacity to explain the sense in which the reasons presented in requests are discretionary. It does so by appealing to the normative role played by discretionary interpersonal valuing-attitudes in grounding those reasons. As such, one way to challenge this account would be to offer an explanation of that discretionary quality without recourse to such interpersonal valuing-attitudes. A rival explanation of exactly this sort is to be found in the approaches by speech act theorists to the phenomena of requesting.

Such approaches tend to follow John Searle (1969, p. 62) in thinking of requesting as expressing a desire of the addressor for the addressee to undertake an action. As such, these approaches concern themselves with what I earlier characterised as a form of epistemic reason-giving. They do not address what I argued in section 1 to be the distinctive normative role of requests as such, wherein a request does something more than merely convey information (either about the addressor's desires, or anything else). But what is more interesting for my purposes is the way in which such pragmaticists think about the discretionary quality of requests in terms of the *indirectness* of requests as speech acts (Searle, 1975). This analytical perspective distinguishes between degrees of (in)directness (Kádár & Haugh, 2013, pp. 23–25). Thus, while a straightforward imperative might be possible, an addressor has the option of deploying layers of indirectness. This could be achieved by phrasing the request as a question ('would you please…'), adding qualifications ('if you wouldn't mind…'), or even merely implying the request by making a related assertion (as when the assertion 'It's a little cold in here' implies the request to close the window). In some instances, an addressor may choose to make a request out of politeness to the addressee. In other instances, this indirectness may function to protect the addressor themselves against embarrassment in the event of the request being refused. Either way, the indirectness of the request is a mechanism by which the addressor communicates her acknowledgement that the reason that she is presenting is discretionary. This is the important point. By focusing on the mechanics of indirect speech acts, one can conceive of the discretionary quality as simply this: an acknowledgement by the addressor, concomitant with the request, that the request creates a reason that the addressee could heed or not, at their discretion.

To be sure, this perspective does look like a challenge to the proposal that I am advancing. The challenge holds that all there is to the discretionary quality of a request is explicable in terms of the communicated acknowledgement of the optional or discretionary force of the reason. This rival explanation threatens to make the idea of the discretionary value of persons superfluous to a theory of requests.

But I do not think that what we have here really is a rival explanation. I do not contest that requests can be made with varying degrees of indirectness. Nor do I contest that such indirectness can, to varying degrees, indicate the addressor's willingness or preparedness to accept the refusal of the request. But the question that I have been addressing in this paper is what normative force, if any, there could possibly be to the discretionary reasons that we create in requests. And to this, as far as I can see, the pragmatic analysis does not propose an answer. In fact, therefore, the pragmatic analysis of discretion begs precisely the question that motivates my proposal.

To see the point here, it may be helpful to attend to the contrast between the kind of reason one can intentionally give another by virtue

28. I am indebted to Basil Vassilicos for raising this challenge and helping me to think it through.

of being authoritative (a command), and the kind one can give without authority (a request). How could the former kind of reason be discretionary? It is difficult to see how a reason could at the same time derive its normative pull from the authority of the addressor, and be discretionary in the sense of decidedly *not obliging* the addressee. Reasons stemming from addressees' authority are — surprisingly enough — authoritative, they are commanding, they are non-obligatory. Having the authority to command someone to Φ does not necessarily entail having the power to give them reasons to Φ with watered-down, less than obligatory strength. As such, the pragmatic analysis, in illustrating the mechanisms through which people present discretionary reasons, thereby illustrates the existence of interpersonally given reasons that do not derive from the authority of the reason-giver. That is the phenomenon targeted by my suggestion for a theory of requests.

Conclusion

I began here by stipulatively defining a request as an attempt by an addressor to create and communicate a non-obligatory reason for the addressee(s) to perform an action. Beyond merely stipulating this definition, I have argued that it captures the conceptual heart of the notion of request, even though it is also fails to match up with the breadth that the term takes on in ordinary usage. On the basis of this definition, I have motivated a general philosophical question about this aspect of our practical lives: what could possibly be the normative grounds of such reasons as those created by requests? In light of the discussion of Enoch's theory of robust reason-giving, of my own proposal, and of related objections, I now have an answer to that question. The normative ground of a successful request, and the reason that the addressee has to accede, is the truth of a prior reason with two conditional elements. That prior conditional reason is of the general form: Person A has a reason to Φ if requested to do so by Person B and if Person A places sufficient discretionary value in Person B. As such, the reasons presented in requests make essential reference to the discretionary value outlooks of their addressees. In particular, they rely on the addressee holding a certain degree of interpersonal valuing-attitude in the person of the addressor. The normativity of requesting, therefore, is a product of a deeper normativity: that value that people have for one another, which can wax and wane in the course of interpersonal interaction.[29]

Works Cited

Blum-Kulka, S., House, J., & Kasper, G. (Eds.). (1989). *Cross-cultural Pragmatics: Requests and Apologies*. New York, NY: Ablex Publishing Corp.

Conant, J., & Rödl, S. (Eds.). (2014). 'The Second Person'. *Philosophical Topics*, 42(1).

Cupit, G. (1994). How Requests (and Promises) Create Obligations. *The Philosophical Quarterly*, 44(177), 439–455.

Dancy, J. (2004). *Ethics Without Principles*. Oxford: Oxford University Press.

Darwall, S. (1977). Two Kinds of Respect. *Ethics*, 88(1), 36–49.

Darwall, S. (2006). *The Second Person Standpoint: Morality, Respect, and Accountability*. Cambridge, MA: Harvard University Press.

Driver, J. (2014). Love and Duty. *Philosophic Exchange*, 44(1), 1–10.

Economidou-Kogetsidis, M. (2010). Cross-Cultural and Situational Variation in Requesting Behaviour: Perceptions of Social Situations and Strategic Usage of Request Patterns. *Journal of Pragmatics: An Interdisciplinary Journal of Language Studies*, 42(8), 2262–2281.

Eilan, N. (Ed.). (2014). 'The Second Person'. *Philosophical Explorations*.

Enoch, D. (2011). Giving Practical Reasons. *Philosophers' Imprint*, 11(4), 1–22.

29. In writing this paper I was helped a great deal by discussions with friends and colleagues. Thanks to audiences at Understanding Value V, and the First Cork Annual Workshop on the Philosophy of Social Agency. And many thanks indeed to those who commented on earlier drafts: Alfred Archer, David Enoch, Alexander Heape, Cristina Roadevin, Bob Stern, Daniel Viehoff, and three anonymous reviewers. The research for this paper was completed as part of a doctoral studentship funded by the White Rose College of the Arts and Humanities, award reference AH/L503848/1.

Huangfu, W. (2012). Cross-Cultural Comparisons of English Request Speech Acts in Native Speakers of English and Chinese. *Cross-Cultural Communication*, 4, 24–29.

Jeske, D. (2008). *Rationality and Moral Theory: How Intimacy Generates Reasons*. New York: Routledge Publishing.

Kádár, D. Z., & Haugh, M. (2013). *Understanding Politeness*. Cambridge: Cambridge University Press.

Kraut, R. (1986). Love De Re. *Midwest Studies in Philosophy*, 10(1), 413–430.

Lance, M., & Kukla, R. (2013). 'Leave the Gun; Take the Cannoli!': The Pragmatic Topography of Second-Personal Calls. *Ethics*, 123(3), 456–478.

Lewis, D. (1989). Dispositional Theories of Value. *Proceedings of the Aristotelian Society, Supplementary Volumes*, 63, 113–137.

Liao, S. M. (2006). The Idea of a Duty to Love. *Journal of Value Inquiry*, 40(1), 1–22.

Markovits, J. (2014). *Moral Reason*. Oxford: Oxford University Press.

Owens, D. (2012). *Shaping the Normative Landscape*. Oxford: Oxford University Press.

Raz, J. (1988). *The Morality of Freedom*. Oxford: Oxford University Press.

Raz, J. (1999). *Practical Reason and Norms*. Oxford: Oxford University Press.

Scheffler, S. (2010). *Equality and Tradition: Questions of Value in Moral and Political Theory*. Oxford: Oxford University Press.

Schroeder, M. (2014). 'Cudworth and Normative Explanations'. In M. Schroeder, *Explaining the Reasons We Share: Explanation and Expression in Ethics* (Vol. 1, pp. 20–41). Oxford: Oxford University Press.

Searle, J. (1969). *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.

Searle, J. (1975). Indirect speech acts. In P. Cole, & J. Morgan (Eds.), *Syntax and Semantics, Vol. 3: Speech Act* (pp. 59–82). New York, NY: Academic Press.

Williams, B. (1981). *Moral Luck: Philosophical Papers 1973–1980*. Cambridge: Cambridge University Press.

Philosophers' Imprint

VOLUME 19, NO. 12

APRIL, 2019

LOGICAL PLURALISM AND LOGICAL NORMATIVITY

Florian Steinberger

Birkbeck College

1. Introduction

This paper explores an apparent tension between two widely held views about logic: that logic is normative and that there are multiple equally legitimate logics. The tension is this. If logic is normative, it tells us something about how we ought to reason. If, as the pluralist would have it, there are several correct logics, those logics make incompatible recommendations as to how we ought to reason. But then which of these logics should we look to for normative guidance? I argue that inasmuch as pluralism draws its motivation from its ability to defuse logical disputes—that is, disputes between advocates of rival logics—it is unable to provide an answer: pluralism collapses into monism with respect to either the strongest or the weakest admissible logic.

The paper proceeds as follows: Section 2 provides a novel analysis of the normative structure of logical disputes. Logical disputes involve various types of normative assessments. In particular, I distinguish *external* assessments that question the correctness of the principles to which the agent assessed holds herself, and *internal* ones by which we criticize the agent for her failure to comply with her own principles. I identify and articulate the principles underlying these normative assessments. Section 3 offers a taxonomy of logical pluralisms and investigates the extent to which each of the taxa leaves room for the aforementioned normative assessments. Section 4 explores the consequences of the fact that an important class of pluralisms—the class that incorporates JC Beall and Greg Restall's influential account—is incompatible with external assessments. I demonstrate that the vulnerability of these views to the well-known 'collapse argument'[1] is a consequence of their inability to account for such assessments. Ultimately

1. See G. Priest, "Logic: One or many?," in J. Woods and B. Brown (eds.), *Logical consequences: Rival approaches* (Oxford: Hermes Scientific Publishers, 2001), G. Priest, *Doubt truth to be a liar* (Oxford: Oxford University Press, 2006), S. Read, "Monism: The one true logic," in D. DeVidi and T. Kenyon (eds.), *A Logical approach to philosophy: Essays in honour of Graham Solomon* (Springer, 2006) and R. Keefe, "What logical pluralism cannot be," *Synthese* 191 (2014).

such forms of pluralism suffer an 'upward' collapse into monism with respect to the strongest admissible logic. Section 5 investigates an alternative form of pluralism according to which logics are correct only relative to their appropriate domains of application. Drawing on the literature on alethic pluralism, I argue that at least when it comes to certain forms of cross-domain discourse such forms of domain-relative pluralism are subject to a different but symmetrically analogous form of 'downward' collapse into monism with respect to the weakest logic. Section 6 argues that on account of the findings of the previous section, the distinction between monism and domain-relative pluralism is merely terminological. Finally, I conclude that the only viable forms of 'pluralism' in light of the normativity of logic are ones that allow for normative conflicts and hence logical rivalry.

Before we proceed a number of preliminary remarks are in order. For one, I rely on the controversial assumption that there is a sense in which logic can be said to be normative. Gilbert Harman has famously challenged the time-honored conception of logic as a normative discipline.[2] His objections have been developed and refined in various interesting ways.[3] I side with those who have sought to rehabilitate the normativity of logic,[4] However, those on the fence about the normative status of logic may read the paper as a conditional claim. Certain kinds of pluralists, who are firmly on the other side of the fence may read it as a *reductio*.[5] What is more, I assume that the connection between principles of logic and norms of reasoning can be rendered explicit in the form of what John McFarlane has called a 'bridge principle.'[6] A bridge principle can be represented schematically as follows:

- ($\star$) If $A_1, \ldots, A_n \models C$, then $N(\alpha(A_1), \ldots, \alpha(A_n), \beta(C))$.

The principle takes the form of material conditional, where the conditional's antecedent states 'facts' about logical consequence and where the principle's consequent sets forth a normative constraint on the agent's doxastic attitudes (belief, disbelief, degree of belief) towards the relevant propositions. The attitudes are represented by 'α' and 'β' on account of the fact that they may be (but need not be) distinct attitudes.[7] Alternatively, a bridge principle's antecedent might appeal not to entailment facts but to the agent's attitudes towards entailments facts:

- ($\star$-γ) If $\gamma(A_1, \ldots, A_n \models C)$, then $N(\alpha(A_1), \ldots, \alpha(A_n), \beta(C))$,

where γ might represent the attitude of knowing, believing, etc. By varying these different parameters, we can generate a considerable number of bridge principles. Here, to illustrate, are three examples:

1. If $A_1, \ldots A_n \models C$, then S ought to believe C, if S believes the A_i.
2. If S believes that $A_1, \ldots A_n \models C$, then S ought not (believe the A_i and disbelieve C).
3. If $A_1, \ldots A_n \models C$, then S has reason to ensure that $cr(C) \geq cr(A_1) + \ldots + cr(A_n) - (n-1)$

In 1. 'ought' takes narrow scope with respect to the conditional in the consequent. It simply states that one's beliefs ought to be closed under logical consequence. 2. is restricted to believed entailment. 'Ought' here takes wide scope over the embedded conditional. Consequently,

2. See G. Harman, *Change in view: Principles of reasoning* (Cambridge: M.I.T. Press, 1986).
3. See *inter alia* S. Dogramaci, "Reverse engineering epistemic rationality," *Philosophy and phenomenological research* 84 (2012), S. Dogramaci, "Communist conventions for deductive reasoning," *Noûs* 49 (2015) and C. Dutilh Novaes, "A dialogical, multi-agent account of the normativity of logic," *Dialectica* 69 (2015).
4. See in particular , J. MacFarlane, In what sense (if any) is logic normative for thought?, 2004, F. Steinberger, "Consequence and normative guidance," *Philosophy and phenomenological research* 95 (2017). See F. Steinberger, "The normative status of logic," in E. Zalta (ed.), *The Stanford encyclopedia of philosophy*, spring 2017 edition (Metaphysics Research Lab, Stanford University, 2017) for a survey of the literature.
5. See C. Blake-Turner and G. Russell, "Logical pluralism without the normativity," *Synthese* (2018) and G. Russell, "Logic isn't normative," *Inquiry* (2017). Note, however, that Beall and Restall (henceforth, 'B&R') declare a firm commitment to the normativity of logic JC Beall and G. Restall, *Logical pluralism* (Oxford: Oxford University Press, 2006).
6. See MacFarlane, *op. cit.*.
7. For simplicity's sake, I set aside suspension of belief.

rather than prescribing a particular belief in the manner of 1., the principle *pro*scribes configurations of attitudes in which the agent simultaneously believes the premises and disbelieves the conclusion. Finally, 3. is a principle governing degrees of belief, represented by the agent's credence function $cr(\cdot)$. Moreover, it employs the defeasible 'has reason'-operator (as opposed to the strict 'ought'). The principle states that the agent has reason to ensure that her degrees of belief respect the stated inequality.[8]

Furthermore, we distinguish three types of normative functions logic might be thought to perform.[9] Logic might be thought to deliver

- *directives*: first-personal instructions guiding the agent in her doxastic conduct;
- *evaluations*: third-personal evaluative standards against which to classify doxastic states as correct or incorrect.
- *appraisals*: third-personal norms that underwrite our attributions of blame and praise to others.

Different bridge principles will be more or less well suited to play a given normative role. For instance, principles like 1. and 3. whose antecedents are insensitive to the agent's recognitional abilities, are unlikely to be serviceable as directives, because ordinary agents with limited logical abilities are in no position to follow them. The same goes for appraisals: it would be inappropriate to fault our epistemic peers for failing to comply with normative principles they cannot possibly live up to. Directives and appraisals may thus be better expressed by attitudinal principles exemplified by 2. That is not to say, however, that there is no use for unrestricted principles; they naturally express objective evaluative standards. After all, the logical coherence of my doxastic state depends on what the logical facts are, not on what I take those facts to be. With these preliminaries in place, let us now turn to our first order of business: the task of analyzing logical disputes.

8. The principle is proposed in this context by H. Field, "What is the normative role of logic?," *Proceedings of the Aristotelian society* 83 (2009b). It is well-known from probability logic, see E. Adams, *A primer of probability logic* (Stanford, CA: CSLI Publications, 1998).
9. See F. Steinberger, "Three ways in which logic might be normative," *Journal of philosophy* (Forthcoming).

2. Logical disputes

The pluralist's role, in logic as elsewhere, tends to consist in defusing disputes she regards as wrongheaded and futile. She seeks to do so by demonstrating how, contrary to appearances, all parties to the dispute can be right. Carnap regarded it as one of the "chief tasks" of *The logical syntax of language* to "eliminate the standpoint" according to which there is but one "correct" logic "together with the pseudo-problems and wearisome controversies which arise as result of it".[10] His deflationary spirit towards logical (and other, in particular metaphysical) disputes is enshrined in his famous principle of tolerance. B&R too reject the very possibility of logical dispute for admissible logics:[11]

> We do not take different logics to be rival analyses of the one fundamental notion (of logical consequence) because we think that the one fundamental notion of logical consequence can be made precise in different ways [...] These different relations are not in competition and they are not rivals. [12]

Logical disputes, then, are ultimately 'based on a confusion' according to the pluralist.[13] Assuming she is right, the pluralist is a heroic character who delivers us from our proclivity for getting embroiled in fruitless squabbles.

With that, let us set the scene. Our story begins prior to the pluralist's appearance, with the logical dispute between Clare and Ira. Clare

10. R. Carnap, *The logical syntax of language* (London: Routledge, 1937), p. xiv
11. Note that B&R do not regard *any* logic as a legitimate contender. Their pluralism is confined to a restricted set of logics that satisfy their admissibility criteria. Inasmuch as B&R and their pluralist brethren seek to dissolve logical disputes with respect to admissible logics, my characterization is nevertheless apt.
12. Beall and Restall, *op. cit.*, p. 88
13. Russell, Logical pluralism, 2014

and Ira are accomplished logicians and philosophers. They agree on a significant number of thorny issues in the philosophy of logic. For instance, both accept my assumption of logic's normativity for reasoning, and that its normative role can be explicitly articulated by means of bridge principles. Also, both are monists: they agree that there is but one correct all-purpose logic. They even agree on what it means for a logic to be correct.[14] But here is the one significant point of disagreement: Clare is an advocate of classical logic while Ira is an advocate of intuitionistic logic. Ira is in the grip of Dummettian arguments in favor of intuitionist revisions of our logical practices; Clare remains unconvinced.[15] Even after countless long nights of well-meaning and intellectually honest debate the two are unable to overcome their differences. It does not matter, for our purposes, who (if either of them) is right. For the sake of the argument, though, let us assume that there is a fact of the matter as to which logic is correct (and that one of the two is).

Let us, then, take a closer look at Clare and Ira's dispute, with a view to making manifest the principles underpinning the normative judgments, assessments and criticisms at the root of their dispute. The principles in question, unlike the closely related standard bridge principles, have not been studied to my knowledge. A proper understanding of the normative structure of logical disputes will thus be of independent philosophical interest. As we will see, though, our analysis has the further benefit of illuminating the normative implications of various forms of pluralism.

In keeping with our provisional assumption of monism, let us begin by spelling out the evaluative standard induced by the correct logic. The following principle captures the idea:

(Objective) If $A_1, \dots A_n \models_{\mathcal{L}} C$, then $N(\alpha(A_1), \dots, \alpha(A_n), \beta(C))$.

(Objective) is a proto-bridge principle. Converting it into a full-fledged principle requires that we specify a good deal of additional information: the type of deontic operator featured, its scope, the type of doxastic attitudes governed, etc. However, even at this level of abstraction, a number of features are noteworthy. For one, the principle's normative role is evaluative. As such, our principle is not in the business of providing direct guidance to the agent, nor does it support criticisms or attributions of blame. Its primary purpose, rather, is to serve as an objective synchronic standard that supports classifications of belief sets into logically 'correct' and logically 'incorrect' ones.[16] Accordingly, assuming that 'ought' is the deontic operator featured in (Objective), 'ought' is itself to be understood as *evaluative*. Unlike *deliberative* or *practical*

14. At a minimum, there are two ways in which logics might be said to be correct, depending on whether one conceives of logic fundamentally as setting forth what we might call (somewhat grandiosely) *the laws of being* or *the laws of thought*. On the former view logic is, much like mathematics, 'about the world' (see e.g. T. Williamson, "Justification, excuses and sceptical scenarios," in F. Dorsch and J. Dutant (eds.), *The new evil demon* (Oxford: Oxford University Press, Forthcoming)). It provides an account of the most general features of reality. On the latter view logic is concerned primarily with our systems of conceptual or linguistic representation. Its aim is, in Priest's words, 'to determine what follows from what—what premises support what conclusions' (Priest, *op. cit.*, p. 196). My distinction is in line with Ole Hjortland's helpful discussion of the opposition of Timothy Williamson's 'deflationary' approach and Graham Priest's 'metalinguistic' approach (O. Hjortland, "Anti-exceptionalism about logic," *Philosophical studies* 174/3 (2016)). My aim here is not to take sides, but simply to note that different conceptions of the nature and purpose of logic entrain different notions of what it means for a logic to be correct. Of course, some philosophers reject the very idea that logics can be meaningfully said to be correct. This, famously, was Carnap's view. It is also endorsed by Field, to whom I return in section 3 (H. Field, "Pluralism in logic," *Review of symbolic logic* 2 (2009c)).

15. I picked the dispute between classical and intuitionistic logic for ease of exposition. Justifiably or not, the Dummettian case for logical revision has somewhat fallen out of fashion. Nothing hangs on the specifics of the case, though. The reader may plug in her favorite argument in support of non-classical logics (quantum logics, relevant logics, dialetheic, paracomplete, supervaluationist, etc.).

16. See Steinberger, *op. cit.* for further discussion. See also e.g. K. Easwaran and B. Fitelson, "Accuracy, coherence, and evidence," in T. Szabo Gendler and J. Hawthorne (eds.), *Oxford studies in epistemology* (2015), fn. 6, 7, 8 and M. Titelbaum, "Rationality's fixed point (Or: In defence of right reason)," in J. Hawthorne (ed.), *Oxford studies in epistemology* (Oxford University Press, 2015), p. 7 for examples of principles of rationality construed as evaluations in this sense.

'ought's, evaluative ones are not agentive. Instead they present certain states of affairs as generally good or correct, and others not. As such it is neither relativized to the agent's ability to recognize entailments (whence the non-relativized antecedent), nor is the 'ought' of the 'can'-implying variety.

For the sake of concreteness, it will be useful in the following to consider a fully articulated principle:

> (Objective -) If $A_1, \ldots A_n \models_{\mathcal{L}} C$, then S ought not (believe the A_i and disbelieve C).

I do not endorse (Objective -) or any other specific principle here. As I discuss in section 3 below, negative principles have certain drawbacks.[17] Nevertheless, it will serve as our go-to principle for purposes of illustration. My aim here is merely to identify the general *form* of the principles underwriting the normative assessments within logical disputes. I leave the task of determining the specifics of the principle's parameter settings for another time.

That said, (Objective -) is *negative* (whence the minus sign), because it enjoins us not to *dis*believe certain propositions (given one's belief in the premises), as opposed to issuing a *positive* injunction to believe. Also, it is a wide scope principle. Finally, '$\mathcal{L}$' stands for whatever logic is in fact correct (in our example, the candidates are the classical consequence relation ($\models_{\mathcal{C}}$) or the intuitionistic one ($\models_{\mathcal{I}}$). The correct logic, whichever it is, induces a corresponding objective evaluative norm.

A surprisingly under-explored fact in the literature on bridge principles is that principles in the mould of (Objective) fail to capture a central dimension of our normative assessments. Let us imagine that intuitionistic logic turns out to be correct. Suppose now that Clare, the classical logician, infers A from $\neg\neg A$ (where she has no independent grounds for believing A).[18] The inference, clearly, falls foul of intuitionistic strictures. However, (Objective) does not tell us this. It detects only *sins of omission*—when an agent fails to appropriately take into account the logical implications of her beliefs; it provides no safeguard against *sins of comission*—when an agent draws inferences that are not sanctioned by the correct logic. What is needed, therefore, is an additional principle that grounds the negative evaluation of Clare's inference:

> (Objective Commissive) Assuming that (Objective) is correct with respect to $\mathcal{L}$ and that S, has no logic-independent grounds for believing C, the following holds:
> If it is the case that (S is permitted to believe C, if S is permitted to believe the A_i), then $A_1, \ldots A_n \models_{\mathcal{L}} C$.

Note that the restriction is indispensable. It might well be permissible for S to believe C on account of C's being non-logically (analytically or materially) entailed by the A_i. Instantiating the principle in the context of our example and contraposing we arrive at the conclusion that since $\neg\neg A \not\models_{\mathcal{I}} A$, Clare's inference to A is impermissible. For symmetry's sake, we should rename (Objective) and it instantiations, '(Objective Omissive)'.

The example of Clare's erroneous inference highlights another important feature of logical disputes: by Clare's own lights, the inference was not erroneous. That is, the correct standards of logical coherence, as codified by (Objective Omissive) and (Objective Commissive), deviate from what Clare takes the correct standards to be: her *subjective evaluative standpoint*. The principle expressing Clare's evaluative standpoint must therefore articulate the evaluative standard to which she holds not just herself, but all of us based on *her* understanding of what the correct consequence relation is. It can be formulated thus:

> (Subjective Omissive) If S endorses $A_1, \ldots A_n \models_{\mathcal{L}} C$, then, S maintains that, for every agent S', $N(\alpha(A_1), \ldots, \alpha(A_n), \beta(C))$.[19]

17. For a fuller discussion, see Steinberger, *op. cit.*.
18. By 'no independent grounds' I mean that there are no grounds for forming the belief in A other than its putative logical relation to $\neg\neg A$. The inference occurs on the basis of Clare's belief or supposition that $\neg\neg A$.
19. Given our assumptions, it is natural to stipulate that one endorses a logic just in case one takes it to be correct or among the best available logics.

The corresponding negative principle is this:

> (Subjective Omissive -) If S endorses $A_1, \ldots A_n \models_{\mathcal{L}} C$, then, S maintains that, for every agent S', S' ought not (believe the A_i and disbelieve C).

Clare and Ira's evaluative standpoints can thus be represented by the appropriate classical and intuitionistic variants of (Subjective Omissive). Both contend that their respective subjective evaluative standpoints are in line with the correct objective evaluative standard represented by (Objective Omissive).

Notice that according to (Subjective Omissive) the agent commits herself to a particular evaluative standard by endorsing a logic, she does not have to endorse particular logical laws for those laws to be normatively binding. That is, in endorsing the logic, the agent willingly takes on a *wholesale* commitment to all concomitant normative demands, whether or not she is in a position to recognize them. (Subjective Omissive), while relativized to the agent, is thus still an evaluative principle and so is not relativized to the agent's beliefs or logical knowledge.

(Subjective Omissive) too requires a commissive counterpart. When Clare infers A from $\neg\neg A$ (where she has no independent grounds for believing A), the inference is licensed from Clare's evaluative standpoint, but not from Ira's thus prompting criticism. Ira's criticism of Clare would thus seem to rely on a subjective version of (Objective Commission):

> (Subjective Commissive) Assuming S endorses (Subjective Omissive) with respect to $\mathcal{L}$ and that the agent assessed, S', has no logic-independent grounds for believing C, the following holds: If it is the case that (S' is permitted to believe C, if S' is permitted to believe the A_i), then $A_1, \ldots A_n \models_{\mathcal{L}} C$.

Imagine now that Ira illicitly (by her own intuitionistic standards) appeals to the law of double negation elimination (DNE) in her reasoning. Clare is well within her rights to criticize Ira. Clearly, though, she does so not because Ira's reasoning is at odds with her (Clare's) classical viewpoint—it patently is not—but because Ira is contravening her own (Ira's) evaluative standards. In other words, Ira manifests a kind of internal incoherence. Call this an *internal* normative assessment in contrast to the principles we have previously encountered all of which underwrite *external* assessments. Internal assessments criticize the agent's failure to reason in conformity with her own evaluative standpoint; external assessments relate to criticisms of the evaluative standard itself. Here is how we might capture internal assessments:

> (Internal Ommissive) S D [endorse $A_1, \ldots A_n \models_{\mathcal{L}} C$ only if P]

Here D is a deontic operator ('ought' or 'has reason') and P is an appropriate pattern of S's attitudes towards the premises A_i and the conclusion C. Spelled out in the manner of our stock example we get:

> (Internal Omissive -) S ought to [endorse $A_1, \ldots A_n \models_{\mathcal{L}} C$ only if (S does not disbelieve C, if S believes the A_i)].

Internal criticisms are grounded in a (presumed) obligation to ensure that one manage one's beliefs in ways consistent with one's own evaluative standards. One is incoherent in this sense if one endorses a logic (and the constraints on belief it imposes) while believing a premise of a valid argument (by that logic's standards) and simultaneously disbelieving its conclusion.

The principle is characterized by its distinctive logical form: the deontic operator takes *super wide scope* over the conditional as a whole (as opposed to familiar wide scope principles that typically operate on the consequent of the main conditional only). S can in principle discharge her obligations in one of two ways: either by conforming to her evaluative standards or by revising those very standards by endorsing a different logic. Though both are live options in principle, in practice the route of logical revision, like that of religious conversion, is one scarcely travelled. One does not renounce one's logical commitments on a whim.

Clearly, as Ira's example shows, internal criticisms might also target

errors of commission relative to the agent's own evaluative standard:

> (Internal Commissive) S D [endorse $\mathcal{L}$ only if (if $A_1, \ldots A_n \models_{\mathcal{L}} C$, then S does not infer C from the A_i, unless there are logic-independent grounds for doing so)].

D, as before, might either be our *ought* or our *has reason*-operator.

A final comment concerning the peculiar normative role performed by internal criticisms: internal criticisms underwrite standpoint-relative third-personal assessments, which makes them appraisals. What makes them peculiar is that while they are relativized to the appraisee's evaluative standpoint, they are not relativized to the appraisee's recognitional capacities. For all we have said, our appraisal of Ira would be equally negative in a case where she fails to take into account an extraordinarily complex intuitionistic argument as it would be in the case in which, in a careless moment, she slips up and helps herself to an application of DNE in everyday reasoning. In both cases she falls foul of her own standards. All the same, we ordinarily do want to distinguish between these two types of failings: the first is due to her all-too-human cognitive limitations; the second is an honest mistake that warrants criticism. To perform both types of assessments, I think of appraisals as variably exigent: on the generous end of the spectrum we allot blame relative to the agent's actual logical capacities; on the unforgiving end of the spectrum we allot blame relative to the agent's evaluative standards regardless of whether she is in a position to live up to them. In between, our appraisals might be relativized to increasingly demanding standards as to which logical implications of the agent's preferred logic she may reasonably be expected to appreciate. (Internal Commissive) sits flatly at the unforgiving end of the spectrum. But it is not hard to see how it might be tempered by restricting the principle's antecedent to the implications the agent takes to obtain or can reasonably be expected to obtain.

In summary, we have uncovered that logical disputes are comprised of two main types of normative assessments: external ones and internal ones. External assessments are concerned with the correctness of the evaluative standards; internal ones are concerned with coherence between the agent's reasoning and the subjective standards to which she holds herself.[20] Among the external assessments we may distinguish (i) the objective evaluative standard and (ii) the agents' subjective evaluative standpoints. Finally, all of these assessments stem from principles, which, in turn, come in two flavors: omissive ones and commissive ones. This rounds up our analysis of the normative structure of logical disputes for now. We return to these principles at the end of the next section. First, though, we must introduce the pluralist.

3. Pluralism

The time has come for the pluralist to make her long awaited appearance. The pluralist maintains that the disputing parties' claims are not genuinely in conflict.[21] For example, although Clare accepts and Ira rejects DNE, the pluralist contends that both can be right. Pluralisms differ over how they account for this possibility. In the following I classify pluralisms accordingly.

Let us immediately set us aside a number of uncontroversial (and hence uninteresting) forms of pluralism. No one doubts that there is a plurality of 'pure logics' in Graham Priest's terminology. [22] Not even the most steadfast monist disputes that there are any number of mathematical structures that we customarily call 'logics' and that may make for worthwhile objects of mathematical study. Nor does the fact that many such logics lend themselves more or less well to different applications—for example, classical propositional logic may be used to model electric circuits, the Lambek calculus naturally models phrase structure grammars, and so on—pose a challenge to the monist. Finally, one may generate a form of pluralism by varying one's logical vocabulary.[23] Which arguments count as valid, depends on our

20. Of course an agent may also criticize peers who share her own (the agent's) evaluative standards when they fail to comply with them. In such cases external and internal criticisms coincide.
21. There are certain exceptions, which we will consider in due course.
22. Priest, *op. cit.*
23. See A. Tarski, "On the concept of logical consequence," in J. Corcoran (ed.),

choices as to which expressions we treat as semantically invariant and which we take to be open to reinterpretation. Russell argues that different conceptions about the nature of the constituents of arguments—i.e. whether we conceive of them as sentences, propositions, statements, etc.—induce different logics.[24] While these accounts certainly make for more interesting forms of pluralism, I nevertheless want to set them aside for present purposes. In what follows I assume that we are working with a fixed set of logical constants and a settled account of the nature of truth-bearers.

Wherein, then, does the disagreement between the monist and pluralist reside? The question of logical pluralism I am after can only be meaningfully raised against the background of the posit that there is, over and above questions of local applicability, a core or 'canonical' [25] application of logic. The pluralist maintains, and the monist disputes, that the core function of logic can be fulfilled by more than one logic. But what exactly does the canonical application of logic amount to? According to Priest, logic's central application is to deductive reasoning. It consists in determining 'what follows from what—what premises support what conclusion—and why' (idem). Philosophers may disagree over the nature of the core application.[26] Regardless of its nature, though, I assume here that there is such a core role for logic to play. A meaningful pluralist challenge amounts to the claim that at least two logics are equally suitable to play the core role.

The first candidate that fits the bill is what sometimes goes by the name of *meaning-variance pluralism*.[27] The label stems from the view's adopted strategy for deflating logical disputes: it is possible for Clare and Ira to both be right because the disputants attach different meanings to the terms involved. Meaning-variance can take multiple forms depending on where the difference in meaning is located. Take the disputed claim that 'The argument form $\ulcorner \neg\neg A \therefore A \urcorner$ is (in)valid'. The semantic difference might be located in the meaning of 'valid', or in the meaning of the logical constants or in both.

Call meaning-variance pluralisms stemming from a difference in the meaning of 'valid' *structural* meaning-variance. A crude version of this view says that 'valid' in Clare's mouth really means 'valid-in-$\mathcal{C}$', whereas in Ira's mouth it means '*in*valid-in-$\mathcal{I}$'. But this misses the point. Of course, no one—classical or intuitionistic logician—has ever disputed *these* claims. The real question is which of the senses of 'valid' (if any) adequately captures *genuine* validity.[28]

A rather more sophisticated brand of structural meaning-variance has been advanced by B&R.[29] According to B&R's influential account, there *is* a core concept of validity, which can be characterized via a set of jointly sufficient and individually necessary conditions—necessary truth-preservation, formality and normativity—and via the so-called

> *Generalized Tarski Thesis*: $\ulcorner \Gamma \therefore_{\mathcal{L}} A \urcorner$ is $valid_{\mathcal{L}}$ if and only if, in every $case_{\mathcal{L}}$ in which all of the members of Γ are true, so is A.

Pluralism arises from the fact that the core concept of validity can be elaborated in several equally legitimate ways depending on how we interpret 'case'.

B&R's structural meaning-variance features prominently in what follows. Yet, to complete the picture, let us briefly turn to the remaining two forms of meaning-variance. *Operational* meaning-variance locates the difference of meaning in (all or some of) the logical connectives.[30]

Logic, semantics, metamathematics, 2 edition (Indianapolis: Hackett, 1983) and A. Varzi, "On logical relativity," *Philosophical issues* 12 (2002).

24. G. Russell, "One true logic?," *Journal of philosophical logic* 37 (2008)

25. Priest, *op. cit.*, p. 196

26. See fn. 14 above.

27. Cf. O. Hjortland, "Logical pluralism, meaning-variance, and verbal disputes," *Australasian journal of philosophy* 91 (2013).

28. One might retort that there is no genuine system-independent concept of validity; that all there is are system-immanent standards of validity. If this were true, we would again be left with a rather uninteresting form of pluralism, not to mention an implausible view of validity.

29. Beall and Restall, *op. cit.*

30. The terminology is inspired by Gentzen-Prawitz-style proof theory, in which inference rules are divided into those that feature specific logical operators (operational rules); and those that codify general constraints on the deducibility relation (structural rules).

On this view, Clare's claim might be understood as '$\neg_{\mathcal{C}}\neg_{\mathcal{C}}A \therefore A$ is valid', whereas Ira's equally correct claim might be read as '$\neg_{\mathcal{I}}\neg_{\mathcal{I}}A \therefore A$ is invalid'. Again, there is no disagreement except, perhaps, over the correct use of the logical connectives.[31]

Finally, on the third view—*hybrid* meaning-variance—the difference resides both in the meaning of 'valid' and in those of the logical operators. Some maintain that structural meaning-variance entails operational meaning-variance.[32] I find it difficult to adjudicate these claims absent a robust account of the meanings of the logical constants. As I am unaware of any such account, I do not pursue this issue further here.

So much for meaning-variance. Let us turn now to a different form of pluralism. Our assumption so far has been that there is what Field has called an 'all-purpose logic'.[33] The assumption enjoys a considerable pedigree. That logic applies unrestrictedly to any subject matter has, in one form or another, been taken to be a non-negotiable component of its job description by many. By contrast, advocates of *domain-relative pluralistm* dispute this characterization.[34] Inquiry, according to them, is irreparably compartmentalized, dividing into several distinct and stable domains.[35] No single logic governs all domains. Rather, different domains call for different logics. And so a logic's normative authority is confined to its proper jurisdiction. We can continue to speak of the canonical application or core role of logic provided we allow for it to be relativized to domains.

Applied to the case of Clare and Ira, the domain-relative pluralist seeks to defuse the dispute by arguing that classical and intuitionistic logic do not compete for the same domain. The dispute is resolved by realizing that both logics have their legitimate domains of application. Of course, this relies on the assumption that the dispute is not domain-internal. And that assumption, it is worth emphasizing, is dubious. After all, the storied conflict between intuitionists and classical logicians has traditionally been a conflict over which of the two logics correctly codifies the standards of correct deductive reasoning in the domain of mathematics. Hence, even if we were to convert Clare and Ira to domain-relative pluralism and they were to agree, for instance, that classical logic governs macroscopic physical objects but that certain observational predicates obey intuitionistic logic, Clare and Ira would still not have made any progress in settling the pivotal question as to which logic to employ in mathematics.

Finally, let us turn to Field's version of logical pluralism. Field's point of the departure is his argument to the effect that 'validity' is not definable in terms of necessary truth-preservation. 'Validity' must be treated as a primitive. Grasping its meaning, however, requires an appreciation of its conceptual role, which, in turn, is characterized by the normative constraints validity imposes on our doxastic attitudes.[36] Field now couples his normative account of validity with his

31. Operational meaning-variance only gives rise to pluralism on the assumption that the alternative meanings are equally legitimate. This is by no means obvious. For example, according to the semantic anti-realist tradition (M. Dummett, *The logical basis of metaphysics* (Cambridge: Harvard University Press, 1991), D. Prawitz, "Meaning and proofs: On the conflict between classical and intuitionistic logic," *Theoria* 43 (1977), N. Tennant, *Anti-realism and logic* (Oxford: Oxford University Press, 1987)) meaning-theoretic considerations reveal the classical meanings of the logical constants to be defective, thus favoring weaker constructive logics.

32. See for instance Priest, *op. cit.*.

33. Field, *op. cit.*, p. 345.

34. O. Bueno and S. Shalkowski, "Modalism and Logical Pluralism," *Mind* 118 (2009)ăand N. da Costa, *Logique classique et non classique: Essaie sur le fondement de la logique* (Paris: Masson, 1997) fall into this category.

35. Domains are typically thought to be individuated by subject matter: think ethics, mathematics, micro-physics, etc. It is worth noting, though, that some phenomena, such as vagueness, cut across domains.

36. Field, *op. cit.*, Field, *op. cit.*, H. Field, "What is logical validity?," in C. Caret and O. Hjortland (eds.), *Foundations of logical consequence* (Oxford University Press, 2015).

non-factualism about the normative.[37] There is, for him, no intelligible sense in which any one set of norms can be said to be uniquely correct.[38] Saying that there is no correct set of logical norms is not say that all logical norms are equally good—some can be better than others. This is because, as a species of epistemic norms, logical norms are selected with a view to promoting our epistemic goals. Logical norms can thus be assessed based on how effectively they achieve this objective. All the same, the picture points to two possible sources of logical pluralism: i) logical pluralism could be a result of pluralism about epistemic goals; ii) even if we agree on the epistemic goals we wish to further, it may be indeterminate which set of norms is most conducive to those goals. We have no reason to assume there to be a unique system that best optimizes for our often competing constraints.

Field's pluralism differs fundamentally from the pluralisms we have encountered so far: Field's pluralism makes room—while the other pluralisms do not—for the possibility of normative conflict.[39] On Field's view there may be multiple competing evaluative standards. It follows that Field's pluralist's does not necessarily fit the mould of the pluralist as a dissolver of logical disputes. On the other hand, Field's picture differs from the standard type of dispute exemplified by Clare and Ira, in that it denies the existence of objective evaluative standards.

This concludes our survey of pluralisms. Let us now marry our findings with those of the previous section by asking which of the normative assessments introduced there have a role to play within the various forms of pluralism. The following table summarizes our findings:

	(Objective)	(Subjective)	(Internal)
Structural MV			✓
Operational MV			✓
Domain-relative	✓-D	✓-D	✓
Non-factualism		✓	✓

By definition, pluralism does away with the notion of a unique correct logic. Consequently, none of our pluralisms allow for an objective bridge principle.[40] The only possible exception is domain-relative pluralism. The way we have portrayed the position it allows for objective, albeit *local* domain-specific bridge principles.[41] What about subjective bridge principles? If there is no correct logic, can I still legitimately take myself and others to be bound by an evaluative standard? Most pluralisms reject this possibility. After all, the point of the pluralist's intervention was to convince us of the futility of logical disputes. Subjective principles have no place within such pluralisms. As before, there are two exceptions. Domain-relative pluralists may countenance local, domain-internal disputes. Also, Field's non-factualism admits of conflicting subjective bridge principles. While there is no fact of the matter as to whether Clare or Ira is right, both may be within their rational rights to adopt and defend their logical policies. Finally, the only type

37. H. Field, "Epistemology without metaphysics," *Philosophical studies* 143 (2009a)

38. This is one of the respects in which Field's pluralism is closer to Carnapian tolerance Carnap, *op. cit.*: both authors explicitly reject the notion that logics can sensibly be called 'correct' or 'true'. If pluralism is narrowly defined as the position that there exist at least two correct logics, their views do not qualify. This goes to show that we should not construe 'logical pluralism' too narrowly.

39. See also T. Kouri Kissel, "Logical pluralism from a pragmatic perspective," *Australasian journal of philosophy* (Forthcoming) and N. Wyatt and G. Payette, "Logical pluralism and logical form," *Logique et analyse* 61 (2018).

40. When we introduced our objective principles, we were working under the provisional assumption of monism. In the present context of neutrality, one may therefore wonder if objective principles must be monist in nature. Consider the modest pluralism of someone who regards classical and intuitionistic logic as equally 'correct'. Instantiating the objective principles above, the likely consequence is, for instance, that one (objectively) may and that one may not infer A from $\neg\neg A$. I confess that I cannot make much sense of such a view (save in the case of epistemic value pluralism, which I discuss below).

41. One could equally imagine a non-factualist variant of the domain-relative pluralism—a hybrid between domain-relative pluralism and Field's non-factualism, if you will—which rejects even local correctness.

of normative assessment that has a place in all pluralist views are internal assessments. In endorsing a logic one commits oneself to the associated norms. Pluralists may allow for me to adopt different logics for different purposes or for particular domains of discourse, but this cannot mean that one gets to pick and choose among the principles of different available logics in the course of one's reasoning as one pleases. Were it permitted to do so, pluralism would collapse into monism with respect to the union of all of the parochial logics. Therefore, given that I operate with a logic within a specified context, I thereby take myself to be bound by the laws of that logic and so am subject to appropriate internal criticisms.

The upshot of these considerations is that meaning variance-based pluralisms are able to accommodate only a very thin, internal normative status. In the next section, I discuss whether this attenuated conception of logical normativity is viable.

4. The collapse argument

My main focus in the following is B&R's pluralism. As we just noted, on B&R's view, a given conception of consequence cannot normatively bind us in virtue of being correct or even by being taken to be so. As far as Clare and Ira's dispute is concerned, neither of their logical practices is susceptible to external criticism. The two have simply elected to play by different, albeit equally acceptable rules. We are left only with a purely system-immanent notion of correctness.

This observation points to a difficulty for B&R's view. Logical norms do not seem to bind us merely in the way that the rules of a game bind us. I take myself to be answerable to the rules of chess only so long as I wish to play chess. Logic, by contrast, is not a game I can choose not to play. Assuming logic is normative for reasoning, its role in our epistemic lives is indispensable. The principles of logic, unlike the rules of a game, are answerable to an external standard, to wit, our broader epistemic aims. Consequently, they must be coordinated with our non-logical epistemic norms.[42]

This, I submit, is the normative source of the so-called 'collapse argument' against B&R's pluralism.[43] The argument, in summary, is this. Suppose that A is known to be true and that B is a (relevant) proposition. Let $\mathcal{L}_1$ and $\mathcal{L}_2$ be two distinct admissible logics such that $\models_{\mathcal{L}_2} \subsetneq \models_{\mathcal{L}_1}$. In particular, suppose that $A \models_{\mathcal{L}_1} B$, but $A \not\models_{\mathcal{L}_2} B$. Do we have logical grounds for believing B? We clearly do on B&R's account. We need not worry that $\mathcal{L}_1$ might lead us astray. After all, $\mathcal{L}_1$ is admissible and so truth-preserving. But if so, the conclusion seems irresistible that, in view of my epistemic aims, I ought to choose an $\mathcal{L}_1$-based bridge principle over the $\mathcal{L}_2$-based principle, lest I pass up the opportunity to come to know B. $\mathcal{L}_1$, as we might put it, normatively dominates $\mathcal{L}_2$. And so one bridge principle—the one featuring the stronger of the two logics—imposes itself, giving rise to the following objective evaluative principle:[44]

> (BP-$\models_{\mathcal{L}_1}$ -) If $\models_{\mathcal{L}_2} \subseteq \models_{\mathcal{L}_1}$, and $A_1, \ldots A_n \models_{\mathcal{L}_1} C$, then S ought not (believe the A_i and disbelieve C).

This suggests that once we factor in our wider epistemic goals, B&R's central claim to the effect that both logics (and their attendant norms) are equally permissible, is false. Notice that the argument does not rely on particularly contentious assumptions about one's epistemic value theory. It merely assumes that, all things being equal, a logic

42. While B&R list normativity among their three admissibility criteria Beall and Restall, *op. cit.*, §2.4, they fail to take the wider epistemic significance of logical normativity into account. Left merely with the internal normative dimension, it is hard to see what work the normativity criterion is doing for them. After all, any consequence relation can trivially be regarded as setting forth norms for anyone who endorses it.

43. See Priest, *op. cit.*, Priest, *op. cit.*, Read, *op. cit.*, Keefe, *op. cit.*, and C. Caret, "The collapse of logical pluralism has been greatly exaggerated," *Erkenntnis* 82 (2017).

44. For simplicity, I present only the variants of our go-to example (Objective Omissive -).

that licences more inferences to potentially epistemically valuable conclusions is to be preferred.

Our conclusion straightforwardly generalizes. Where we are confronted with various admissible logics that are totally ordered in terms of strength, we simply pick the strongest of the bunch.[45] In cases where the admissible logics are not totally ordered, the lesson that we ought to exploit our logical resources still applies. In the simplest case where we have two admissible logics, $\mathcal{L}_1$ and $\mathcal{L}_2$, that are incomparable with respect to inclusion (as, for example, in the case of intuitionistic logic and a standard system of relevant logic), the apposite principle would seem to be:

(BP-($\models_{\mathcal{L}_1} \vee \models_{\mathcal{L}_2}$)) If $A_1, \ldots A_n \models_{\mathcal{L}_1} C$ or $A_1, \ldots A_n \models_{\mathcal{L}_2} C$, then S ought not (believe the A_i and not believe C).

Generalizing beyond the case of two logics, we arrive at the following:

(BP-$\cup \models_{\mathcal{L}_i}$) If there exists an admissible $\mathcal{L}_i$, such that $A_1, \ldots A_n \models_{\mathcal{L}_i} C$, then S ought not (believe the A_i and not believe C).[46]

The upshot of these reflections is that B&R's pluralism is vulnerable to a kind of *upward* collapse. Once our broader epistemic commitments are duly taken into account, it looks as if we ought to adopt the strongest available consequence relation among our admissible logics.[47] We thus find ourselves bereft of any rationale for endorsing a weaker logic.

How might B&R respond to the collapse worry? B&R are advocates of negative bridge principles: 'if an argument is valid,' they write, 'then you somehow go *wrong* if you accept the premises but reject the conclusion'.[48] The collapse argument is driven by the fact that in opting for a weaker logic one forgoes the opportunity to acquire an epistemically valuable belief. Perhaps, though, it is a mistake to construe the normativity of logic as issuing obligations to believe—even wide-scope ones. B&R's favored negative bridge principles are mere safeguards of logical coherence: I can comply with the bridge principle, simply by not bearing any kind of attitude at all towards the conclusion of a valid argument, just so long as I do not 'actively' disbelieve it (while believing the premises). However, this response is of little help to the pluralist, even if negative principles were to win the day. For even according to our negative principle the weaker logic $\mathcal{L}_2$ fares worse epistemically than $\mathcal{L}_1$: plainly, $\mathcal{L}_2$ permits disbelieving true propositions (and $\mathcal{L}_1$-consequences) such as B.[49]

Colin Caret has proposed a different response on behalf of B&R.[50] Following Hjortland and Shapiro, Caret proposes to interpret B&R's version of meaning-variance as a form of contextualism about the meaning of the validity predicate.[51] The predicate's meaning must be understood relative to a contextually determined standard of logical strictness. Certain types of cases (incomplete ones, inconsistent ones, etc.) will be live options in some contexts, thus raising the strictness bar by requiring us to consider a larger class of cases; other contexts will impose laxer standards allowing us to disregard certain cases thus

45. As before, I am assuming that logics are ordered by inclusion over their consequence relations.
46. A word of caution is in order here. The taking of unions of consequence relations may result in a trivial system. An example is given by Abelian and classical propositional logic Read, *op. cit.*: in Abelian logic we have $\neg A, B \models_{\mathcal{A}} ((A \to B) \to B) \to A$, whereas classical logic yields $\neg A, B \models_{\mathcal{C}} \neg(((A \to B) \to B) \to A)$.
47. Cf. Keefe, *op. cit.*.
48. Beall and Restall, *op. cit.*, p.16.
49. What is more, as MacFarlane, *op. cit.* remarks, negative principles seem too weak, at least on their own. Take, for instance, the case of the aforementioned (Subjective Omissive -). Suppose my colleague refutes the claim $A \wedge B$. She rightly points out that I have previously professed belief in both A and B, though separately. Intuitively, I am under rational pressure to abandon at least one of my beliefs. But the negative principle does not account for that pressure. Instead it affords me a dubious loophole: my endorsing A and B merely provides me with an obligation *not to disbelieve* $A \wedge B$. Surely, though, the situation demands more of me. It demands that I own up to my doxastic commitment towards $A \wedge B$.
50. Caret, *op. cit.*
51. See Hjortland, *op. cit.* and S. Shapiro, *Varieties of logic* (Oxford: Oxford University Press, 2014).

leaving room for 'more' logical implications. Since strict contexts mandate weaker logics, contextualism appears to stave off the threat of collapse by providing the previously lacking rationale for espousing a weaker logic.

Caret's idea is elegant, but not ultimately convincing in my view. For one, the notion of a variable standard of logical strictness lacks motivation. Caret models his proposal after epistemic contextualism, according to (a version of) which, 'knows' is to be interpreted relative to a contextual parameter expressing an epistemic standard. Different contexts call for different standards, thus altering the extension of 'knows' accordingly. Wherever one ultimately stands on the viability of epistemic contextualism, it is hard to deny that the position enjoys at least a *prima facie* intuitive pull. The same cannot be said for Caret's proposal. I see no good reason for thinking that our validity judgments are in fact sensitive to a strictness parameter, nor, for that matter, that they should be. We simply do not ordinarily recognize contexts that select for different stricter or laxer logical standards and so for weaker and stronger logics. To be sure, we may at times 'try on' different logics as a possible way of resolving a paradox (as in Caret's example of the liar (idem)) or to accommodate persistently recalcitrant data. But such cases are more readily thought of as instances of suppositional reasoning. In much the same way in which I might posit the truth of certain propositions to explore their consequences in the course of theoretical deliberation, I might posit the validity or invalidity of a principle of logic in order to weigh the costs and benefits of each of my options, e.g. 'Ought I to restrict the truth-predicate or should I revise my logic in order to account for semantic paradoxes?'[52] But engaging in deliberation of this kind does not commit me to logical pluralism, nor is there any need to wheel in contextualist machinery to make sense of it.

52. Peter P. Schroeder-Heister, "A natural extension of natural deduction," *Journal of symbolic logic* 49 (1984) proposes a natural deduction calculus that allows for the introduction of dischargeable deductive rules in the context of suppositions, which can be thought of as a proof-theoretic representation of such logical suppositions.

Let us turn now to the final response. The collapse argument makes certain assumptions about the role of logic in our epistemic lives and about the epistemic ends we pursue. These assumptions are sufficiently weak to be compatible with a wide variety of epistemic value theories. It might, however, be said to rely on the tacit assumption of epistemic monism in that it presumed there to be but one fundamental epistemic value (e.g. truth or knowledge). Perhaps, then, the absent motive for logical pluralism resides in pluralism about epistemic value. Weaker logics impose constraints beyond mere truth-preservation on the notion of logical consequence. Perhaps some of these constraints can be motivated by appeal to alternative epistemic values? Perhaps so. But even if value pluralism is correct, we are still owed a story as to what these values are and how the candidate logics might promote or respect these values. As things stand, it is hard to make out even the contours of such an account. Certainly, it is unclear how these logics might relate to the types of values often invoked such as understanding or wisdom.

5. Domain-relative pluralism

Let us then consider a different strategy for dodging the collapse worry. Since this strategy amounts to a more radical departure from B&R's pluralism, it merits separate treatment. The strategy consists in parrying the collapse argument by espousing domain-relative pluralism.[53] Domain-relative pluralism, recall, is the view that different domains of discourse select different logics. The domain-relativist's response to the collapse argument is simple: while the stronger logic in the example above may be appropriate for some domains, some domains may not support all of its implications and so may require a weaker logic. The

53. B&R are unequivocal in their rejection of domain-relative pluralism, see Beall and Restall, *op. cit.*, p. 88.

threat of collapse is thus blocked by compartmentalizing our logics' domains of application.

Domain-relative pluralism raises an important question well known in the literature on alethic pluralism as the *problem of mixed compounds*. Adapted to our present concerns, the question can be put thus: What are we to make of logically complex propositions the atomic parts of which pertain to domains governed by distinct logics? And what are we to make of inferences involving premises pertaining to domains governed by distinct logics? Far from being a niche phenomenon, cross-domain reasoning is commonplace and of central importance to our intellectual pursuits. The mathematical, the physical, the ethical, the legal, the aesthetic and so on are frequently intermingled in our attempts to make sense of the world. Domain-relative pluralists must therefore be able to account for propositions and inferences that straddle muliple domains.

How might the pluralist approach this challenge? To keep things simple, consider a toy example involving just two domains: that of mathematics, D_M, and that of macro-physics, D_P. In keeping with our story line, let us assume that our pluralist endorses intuitionistic logic within the mathematical domain and classical logic within the physical domain. Now let A be a mathematical proposition and B a physical proposition, both true in their respective domains. Given these assumptions, the question is this: What are we to make of $A \star B$, where $\star$ is a logical connective? There are three possibilities:

- Treat $A \star B$ as if it belonged to D_M;
- Treat $A \star B$ as if it belonged to D_P;
- Treat $A \star B$ as belonging to $D_M \bullet D_P$,

where $D_M \bullet D_P$ is a status that functionally depends on D_M and D_P but is distinct from both.

Nikolaj Pedersen and Cory Wright, in their structurally analogous discussion of alethic pluralism, go in for the first option.[54] It will be helpful to introduce some terminology. Let us again assume a partial ordering, $\leq$, by inclusion over our logics. Let us say that, for any proposition P, $\lambda(P)$ is the logic governing P in virtue of the domain to which P pertains. In our example, we have $\lambda(A) = \mathcal{I}$ and $\lambda(B) = \mathcal{C}$ and thus $\lambda(A) \leq \lambda(B)$. Following a standard move in algebraic semantics, Pederson and Wright now treat conjunction and disjunction as 'minimizing' and 'maximizing' operations respectively. In our context this amounts to:

$$\lambda(A \wedge B) = min(\lambda(A), \lambda(B))$$
$$\lambda(A \vee B) = max(\lambda(A), \lambda(B))$$

But this cannot be quite right as the following simple argument reveals. Suppose I prove $\neg\neg C$, for some $C \in D_M$, where C is not effectively decidable. Because D_M is governed by intuitionistic logic, I am not permitted to infer C. However, if the proposal were correct, I would have a ready-made strategy for circumventing these intuitionistic strictures. Simply disjoin the conclusion with a random physical falsehood, P, yielding $\neg\neg C \vee P$. But $\neg\neg C \vee P$ is subject to classical logic and so is equivalent to $C \vee P$. And since we know that $\neg P$, an application of disjunctive syllogism yields the purely mathematical C. Nothing hangs

54. See N. J. L. L. Pedersen and C.D. Wright, "Pluralist theories of truth," in E. Zalta (ed.), *Stanford encyclodedia of philosophy* (2013), §4.5.2. The objection to follow does not necessarily apply to them. It does, however, carry over if the different truth-properties associated with the two domains were to induce different logics. While alethic pluralism certainly does not entail logical pluralism, the former does naturally entrain the latter given certain assumptions, see M. Lynch, *Truth as one and many* (Oxford: Oxford University Press, 2009), p. 95–96 and N. J. L. L. Pedersen, "Pluralism x 3: Truth, logic, metaphysics," *Erkenntnis* 79 (2014) himself. Going in the opposite direction, S. Read, "Review of J.C. Beall and G. Restall, *Logical pluralism*, Clarendon Press, 2006," *Notre Dame philosophical reviews* (2006) argues that B&R would do well to endorse alethic pluralism. It should be emphasized, though, that B&R reject alethic pluralism (p. 100).

on the specifics of my example. The same (or analogous) arguments can be generated for similar cases.

To avoid such difficulties a retreat to what Lynch has called 'logical modesty' recommends itself:

> **Logical modesty:** Where a compound proposition or inference contains propositions from distinct domains, the default governing logic is that of the compound or inference's weakest member.[55]

Logical modesty is a plausible stance. However, as it stands it presupposes the comparability of all logics involved, which, we said, is not always possible. For instance, we have $\models_{\mathcal{I}} \not\subseteq \models_{\mathcal{R}}$ and $\models_{\mathcal{R}} \not\subseteq \models_{\mathcal{I}}$ (where $\mathcal{I}$ is intuitionistic logic and $\mathcal{R}$ the system of relevant implication.) How to proceed? In analogy with our development of the collapse argument in the previous section where we took the union of the relevant consequence relations, the natural move here is instead to take the intersection of the logics in question. This is in the spirit of logical modesty: When engaging in cross-domain reasoning, we should draw only on principles sanctioned by *all* the relevant logics. We thus arrive at the following

- (BP-$\cap$ $\models_{\mathcal{L}_i}$) If for all $\mathcal{L}_i$, $A \models_{\mathcal{L}_i} C$, then S ought not (believe A and disbelieve C).

Thus, whereas the collapse argument results in an *upward* collapse into monism, domain-relative pluralism gives rise to a *downward* collapse. The direction of the collapse is determined by whether the admissible logics are reliable or not. In the context of B&R's pluralism all logics in question are admissible and so necessarily truth-preserving. In the present case, different logics can be reliably applied only in their appropriate domains. Misapplying a stronger logic in a 'weaker' domain may lead us from truth to falsity.

The question now is whether the downward collapse of domain-relative pluralism when it comes to cross-domain discourse also amounts to an all-out collapse into *monism*? I turn to this question in the next section.

55. Lynch, *op. cit.*, p. 100.

6. Duck-rabbit pluralism

It will not have escaped the attentive reader's attention that Clare and Ira have been absent throughout our discussion of pluralism.[56] Happily, they are making a reappearance in our present discussion. It is important to bear in mind for the upcoming act that, their obvious differences of opinion notwithstanding, Clare and Ira's philosophies of logic are largely aligned. In particular, both are staunch monists.

Let us focus on Ira. Ira, as we know, rejects classical logic in favor of intuitionistic logic. She does so because she maintains that characteristically classical principles lack universal validity and so cannot form part of the correct logic. Intuitionistic principles, by contrast, do hold without fail in all domains according to her. Ira's view does not prevent her from calling upon classical principles when reasoning in circumstances in which she thinks they do hold. However, in so doing she accords classical principles the status of domain-specific *non-logical* principles of inference, much in the way in which one might legitimately appeal to the principle that the whole is greater than its proper parts outside of infinitary set theory.[57]

In a dramatic twist, a third character steps on the scene: Dora. Dora agrees with Ira both *that* intuitionistic restrictions of classical logic are warranted and *where* these are warranted. The twist, though, is that Dora is a domain-relative pluralist. Where for Ira a logical principle's membership in the correct logic and its universal validity are necessarily linked, Dora's position is that the two may come apart. That is, where Ira views local failures of validity as decisive demonstrations that classical principles have no place in the correct logic, Dora does

56. I am borrowing the phrase 'duck-rabbit' pluralism from Priest, *op. cit.*.

57. I set aside recent accounts of infinite sets that preserve the part-whole principle. See P. Mancosu, "Measuring the size of infinite collections of natural numbers: Was Cantor's theory of infinite number inevitable?," *The review of symbolic logic* 2 (2009) for discussion.

treat classical principles as genuinely valid, albeit only within the confines of their rightful domain.

Hence, while Ira and Dora deploy the same principles in the same contexts, the difference between their views stems from their conflicting verdicts regarding the logical status of classical principles. For example, both condone the use of DNE when reasoning about decidable domains. However, Dora treats the principle as a logical validity *relative to the appropriate domains*, whereas Ira insists on treating it as non-logical because merely domain-specific. The two are thus in complete agreement with respect to the norms of reasoning to which they hold themselves. They disagree only in that Dora uses 'valid', 'logical' and their cognates more liberally, while Ira reserves these honorifics for principles she takes to be universally applicable. Viewed in this way, the conflict between the monist and the domain-relative pluralist seems to come down to a mere semantic squabble. After all, what substantive questions could possibly hang on our being more or less liberal in our application of 'valid' and 'logical principle'?[58]

One might be tempted to point to the fact that Ira's stricter interpretation has the longstanding tradition of treating its *formality* or 'lack of subject matter' as partially definitive of logic on its side.[59] What characterizes logic as a discipline (at least in part), on this view, is its unrestricted applicability. But these considerations are of little succor to Ira. After all, we already knew that Ira does, while Dora does not, build universal applicability into her conception of logicality. The question is whether there are good reasons for doing so. And the trouble is that accounts of logicality in this tradition do not deliver on reasons. Logic's universality, rather, is posited as an unexplained explainer.

A *prima facie* more promising objection has been levelled at the monist.[60] By virtue of her uncompromising conception of logicality, the monist effectively adopts a position of logical modesty with respect to all discourse (not merely for cross-domain discourse). According to her, the only *bona fide* laws of logic are those that hold good in all domains. But here's the rub: scarcely any logical principle has gone unchallenged in one context or another. Hence, if for sufficiently many domains our best overall theory requires weakening our logic, the monist runs the risk of finding herself with an unworkably weak or even empty consequence relation. Call this the *Objection From the Threat of Logical Nihilism*.

The threat of nihilism also seems to show Dora to be in an advantageous position when making potentially logic-altering theoretical decisions. For let us suppose our best theory of a given domain is faced with persistent recalcitrant data. Let us assume, moreover, that we could accommodate the data by either revising the theory in question or by locally abandoning certain logical laws. How would Ira and Dora approach this theoretical choice? Ira, the monist, would seem to have a very strong incentive not to tinker with her logic lest she ends up with a cripplingly weak all-purpose logic. These global theoretical considerations thus impose stiff constraints on Ira's local theoretical choices. Dora, by contrast, appears to enjoy a great deal more flexibility, which would appear to be an asset.

But this picture is misleading. The trouble is that it overlooks the insights from our discussion of the downward collapse problem. For when it comes to cross-domain discourse, the pluralist and the monist are in the same boat: both are equally committed to logical modesty when several domains are involved. In such cases both must make do with the principles that hold in all the relevant domains. It follows that, local logical revision in response to theoretical pressures are likely to come at a heavy cost also for the pluralist. Ira and Dora are thus both subject to a standing *pro tanto* injunction in favor of logical conservativeness.

True, the pluralist's and and monist's dialectical situations are not identical: while the monist is *always* committed to a core logic applicable across all domains, the pluralist must resort to logical modesty only

58. This discussion draws in part on Priest, *op. cit.*, p. 203
59. J. MacFarlane, *What does it mean to say that logic is formal?* (Ph. D. diss.), Pittsburgh: University of Pittsburgh, 2000
60. B&R themselves advance this argument. See Beall and Restall, *op. cit.*, p. 93. See also Bueno and Shalkowski, *op. cit.*.

in 'worst-case' scenarios involving particular restrictive cross-domain discourse. But the force of this response is significantly mitigated by the following two considerations. First, recall that the monist too can augment her core logic by introducing domain-specific (non-logical principles) where appropriate. As we have seen in the case of Ira and Dora, both the monist and domain-relative pluralist have access to the same principles. Simply, Ira is more sparing in which of the principles she considers genuinely logical. Second, cross-domain discourse, far from being a fringe phenomenon, is crucial to our intellectual pursuits. As Lynch aptly puts it

> reason, by its nature, is universal in its scope—it allows us to combine propositions from different domains into more complex propositions, and to make inferences across different subjects—as when we draw moral conclusions from partly non-moral premises.[61]

In short, the threat of nihilism (or at least the threat of an impractically weak logic) afflicts both the domain-relative pluralist and the monist to a significant degree. It is not clear, to say the least, that this tips the balance in the pluralist's favor.

Is there anything, then, that could convince us that Ira and Dora are embroiled in more than a terminological tangle? One option would be for Ira to show that there is more to (what she calls) *genuine* logical principles than universal validity—some distinctive property that would set properly logical principles apart from merely domain-specific principles. Different types of accounts are conceivable: genuine logicality might manifest itself by way of distinct metaphysical property or perhaps via a distinctive normative profile. I explore neither of these options here. If either could be shown to stick, this would demonstrate the illegitimacy of the domain-relativist's description of the situation. But even in the absence of such a demonstration, our discussion has shown that domain-relative pluralism amounts to nothing more than a re-description of monism. We have found no good reason for choosing the pluralist duck over the monist rabbit.

61. Lynch, *op. cit.*, p. 86

7. Conclusion

Here is what we have established. We have analysed the normative structure of logical disputes and we have provided a classification of logical pluralist views in accordance with their strategy for resolving such disputes. Among the forms of pluralism that offered such a strategy at all, we distinguished meaning-variance pluralisms and (certain) domain-relative pluralisms. The former leave no room for external assessments, the latter allow for external assessments when it comes to disputes about a particular domain. I argued that both types of pluralism (or at least the viable representatives thereof) ultimately collapse into monism. Consequently, the only forms of genuine logical pluralism compatible with the normativity of logic are ones that allow for logical disputes. More succinctly put: if logic is normative, competition between logics may be inevitable.

Acknowledgments I am particularly grateful to Filippo Ferrari, Matthias Rolffs, Dorothee Schmidt and Erik Stei whose kind invitation to the Pluralisms Workshop in Bonn in March 2017 prompted me to write this paper. I would like to thank the audiences in Bonn and at the Proofs Workshop at the Institute of History and Philosophy of Science and Technology in Paris, as well as my colleagues and students at Birkbeck's Work in Progress Seminar and the Philosophy Society for valuable feedback. Finally, I would like to thank two anonymous referees for their helpful suggestions.

References

Adams, E. *A primer of probability logic*. Stanford, CA: CSLI Publications, 1998

Beall, JC and G. Restall *Logical pluralism*. Oxford: Oxford University Press, 2006

Blake-Turner, C. and G. Russell "Logical pluralism without the norma-

tivity" *Synthese* (2018)
Bueno, O. and S. Shalkowski "Modalism and Logical Pluralism" *Mind* 118 (2009): 295–321
Caret, C. "The collapse of logical pluralism has been greatly exaggerated" *Erkenntnis* 82 (2017): 739–760
Carnap, R. *The logical syntax of language*. London: Routledge, 1937
da Costa, N. *Logique classique et non classique: Essaie sur le fondement de la logique*. Paris: Masson, 1997
Dogramaci, S. "Reverse engineering epistemic rationality" *Philosophy and phenomenological research* 84 (2012): 513–530
_____ "Communist conventions for deductive reasoning" *Noûs* 49 (2015): 776–799
Dummett, M. *The logical basis of metaphysics*. Cambridge: Harvard University Press, 1991
Dutilh Novaes, C. "A dialogical, multi-agent account of the normativity of logic" *Dialectica* 69 (2015): 587–609
Easwaran, K. and B. Fitelson "Accuracy, coherence, and evidence" In T. Szabo Gendler and J. Hawthorne (eds.), *Oxford studies in epistemology*. 2015: 61–96
Field, H. "Epistemology without metaphysics" *Philosophical studies* 143 (2009a): 249–290
_____ "What is the normative role of logic?" *Proceedings of the Aristotelian society* 83 (2009b): 251–268
_____ "Pluralism in logic" *Review of symbolic logic* 2 (2009c): 342–359
_____ "What is logical validity?" In C. Caret and O. Hjortland (eds.), *Foundations of logical consequence*. Oxford University Press, 2015
Harman, G. *Change in view: Principles of reasoning*. Cambridge: M.I.T. Press, 1986
Hjortland, O. "Logical pluralism, meaning-variance, and verbal disputes" *Australasian journal of philosophy* 91 (2013): 355–373
_____ "Anti-exceptionalism about logic" *Philosophical studies* 174/3 (2016): 631–658
Keefe, R. "What logical pluralism cannot be" *Synthese* 191 (2014): 1375–1390
Kouri Kissel, T. "Logical pluralism from a pragmatic perspective" *Australasian journal of philosophy* (Forthcoming)
Lynch, M. *Truth as one and many*. Oxford: Oxford University Press, 2009
MacFarlane, J. *What does it mean to say that logic is formal?*. Ph. D. diss Pittsburgh: University of Pittsburgh 2000
_____ In what sense (if any) is logic normative for thought? 2004 Unpublished
Mancosu, P. "Measuring the size of infinite collections of natural numbers: Was Cantor's theory of infinite number inevitable?" *The review of symbolic logic* 2 (2009): 612–646
Pedersen, N. J. L. L. "Pluralism x 3: Truth, logic, metaphysics" *Erkenntnis* 79 (2014): 259–277
Pedersen, N. J. L. L. and C.D. Wright "Pluralist theories of truth" In E. Zalta (ed.), *Stanford encyclodedia of philosophy*. 2013
Prawitz, D. "Meaning and proofs: On the conflict between classical and intuitionistic logic" *Theoria* 43 (1977): 2–40
Priest, G. "Logic: One or many?" In J. Woods and B. Brown (eds.), *Logical consequences: Rival approaches*. Oxford: Hermes Scientific Publishers, 2001: 23–38
_____ *Doubt truth to be a liar*. Oxford: Oxford University Press, 2006
Read, S. "Monism: The one true logic" In D. DeVidi and T. Kenyon (eds.), *A Logical approach to philosophy: Essays in honour of Graham Solomon*. Springer, 2006: 193–209
_____ "Review of J.C. Beall and G. Restall, *Logical pluralism*, Clarendon Press, 2006" *Notre Dame philosophical reviews* (2006)
Russell, Logical pluralism 2014
Russell, G. "One true logic?" *Journal of philosophical logic* 37 (2008): 593–611
_____ "Logic isn't normative" *Inquiry* (2017)
Schroeder-Heister, P. "A natural extension of natural deduction" *Journal of symbolic logic* 49 (1984): 1284–1300
Shapiro, S. *Varieties of logic*. Oxford: Oxford University Press, 2014
Steinberger, F. "Consequence and normative guidance" *Philosophy and phenomenological research* 95 (2017)

_____ "The normative status of logic" In E. Zalta (ed.), *The Stanford encyclopedia of philosophy* Spring 2017 edition. Metaphysics Research Lab, Stanford University, 2017

_____ "Three ways in which logic might be normative" *Journal of philosophy* (Forthcoming)

Tarski, A. "On the concept of logical consequence" In J. Corcoran (ed.), *Logic, semantics, metamathematics* 2 edition. Indianapolis: Hackett, 1983: 409–420

Tennant, N. *Anti-realism and logic*. Oxford: Oxford University Press, 1987

Titelbaum, M. "Rationality's fixed point (Or: In defence of right reason)" In J. Hawthorne (ed.), *Oxford studies in epistemology*. Oxford University Press, 2015

Varzi, A. "On logical relativity" *Philosophical issues* 12 (2002): 197–219

Williamson, T. "Justification, excuses and sceptical scenarios" In F. Dorsch and J. Dutant (eds.), *The new evil demon*. Oxford: Oxford University Press, Forthcoming

Wyatt, N. and G. Payette "Logical pluralism and logical form" *Logique et analyse* 61 (2018): 25–42

METAPHILOSOPHY

METAPHILOSOPHY
Vol. 50, Nos. 1–2, January 2019
0026-1068

NORMATIVITY IN THE PHILOSOPHY OF SCIENCE

MARIE I. KAISER

Abstract: This paper analyzes what it means for philosophy of science to be normative. It argues that normativity is a multifaceted phenomenon rather than a general feature that a philosophical theory either has or lacks. It analyzes the normativity of philosophy of science by articulating three ways in which a philosophical theory can be normative. Methodological normativity arises from normative assumptions that philosophers make when they select, interpret, evaluate, and mutually adjust relevant empirical information, on which they base their philosophical theories. Object normativity emerges from the fact that the object of philosophical theorizing can itself be normative, such as when philosophers discuss epistemic norms in science. Metanormativity arises from the kind of claims that a philosophical theory contains, such as normative claims about science as it should be. Distinguishing these three kinds of normativity gives rise to a nuanced and illuminating view of how philosophy of science can be normative.

Keywords: normative, scientific practice, practice turn, philosophical methods, empirical information, epistemic norms.

1. Introduction

In recent decades, many areas of philosophy of science have undergone what is now referred to as a "practice turn" (e.g., Soler et al. 2014), that is, a turn towards scientific practice. More and more philosophers of science agree that philosophical theories about science must account for how science actually is done and must be informed, for instance, by the explanations developed in scientific practice and by the investigative strategies that scientists in fact employ. In other words, they agree that philosophical accounts about science must arise from an "empirical engagement with science" (Boumans and Leonelli 2013, 260) and that philosophers of science should seek to understand—"from the inside while retaining a philosophical perspective" (Wimsatt 2007, 27)—how science works, why it is successful, and why it sometimes fails. Philosophy of science that pays close attention to scientific practice is also called

"philosophy of science in practice" (Ankeny et al. 2011; Mansnerus and Wagenknecht 2015).[1]

Philosophers spell out in different ways what it means for philosophy of science to turn towards scientific practice. Frequently, the practice turn is associated with a shift from analyzing products of science to analyzing scientific processes or activities (Chang 2011), with a shift from theory-focused to practice-centered epistemology (Waters 2014), and with a shift from purely intellectual and conceptual perspectives on science to ones that also consider the social context and the material aspects of scientific practice (Soler et al. 2014). Moreover, the turn towards scientific practice is simultaneously understood as a turn away from traditional normative theories about science, which construct ideals of how science should or ideally would work. These normative ideals were typically formulated ex cathedra and are thus criticized for being disconnected from and peripheral to the empirical reality of scientific practice. Accordingly, Soler and colleagues characterize the practice turn as a "shift from normative to descriptive perspectives on science" (2014, 15).

On the other hand, normativity continues to play a role in the philosophy of science after the practice turn. Several philosophers who pay close attention to scientific practice emphasize that science is an inherently collective activity and that we thus must take into account the social norms that influence the pursuit of scientific knowledge (Lloyd 2006; Kitcher 2011). In addition, some philosophers subscribe to the practice turn but sustain their normative aspirations. These philosophers use the results of their descriptive analyses to offer normative advice about how science should be done and about how certain concepts should be understood (e.g., Woodward 2003, 7; Craver 2007, viii). Finally, even if philosophers of science seek to describe a certain element of scientific practice, they seem to implicitly rely on normative assumptions, including assumptions about what are good examples and about how to assess the success of science. Hence, philosophy of science in practice seems to be thoroughly normative.

How can it be that the practice turn involves both moving away from normativity and at the same time leaving room for and even moving towards normativity? My goal in this paper is to solve this apparent contradiction. I argue that normativity in the philosophy of science is not a single matter, not a general feature that a philosophical account either has or doesn't have, but is a multifaceted phenomenon. The turn towards scientific practice involves moving away from one kind of normativity but not from others. I analyze in the paper the normativity of philosophy of science by articulating three ways in which a philosophical account can be normative. I distinguish metanormativity, methodological normativity,

[1] A recent indication of the turn towards scientific practice is the Society for Philosophy of Science in Practice (SPSP), founded in 2005.

and object normativity, and show how the different kinds of normativity relate to each other.

The structure of the paper is as follows. In section 2, I introduce the terminological framework of my analysis. Section 3 explicates the first way in which philosophy of science can be normative, which I call "metanormativity." I show that this type of normativity emerges from the kind of claims that a philosophical theory about some feature or element of science contains. I argue that philosophers of science who join the practice turn move away from metanormativity in the ex cathedra style, but that they need not abandon metanormativity in general. In Section 4, I examine the methodology of practice-oriented philosophy of science and reveal a second kind of normativity. My central claim is that methodological normativity arises from the need to make normative assumptions in selecting, interpreting, and evaluating the empirical basis of a philosophical account. In Section 5, I distinguish a third kind of normativity and argue that it arises from the fact that some philosophers of science discuss the role of epistemic and social norms in science. I refer to this kind of normativity as "object normativity" because it is due to the object of philosophical theorizing itself being normative. Finally, in section 6, I point out how the three kinds of normativity can be combined and how they depend on each other.

2. Terminological Framework

This section introduces the basic concepts on which my metaphilosophical analysis relies as well as the figure, to be successively refined in subsequent sections, that I use to illustrate my claims. The philosophy of science is a philosophical discipline that can be said to consist of different philosophical accounts or theories. You might prefer to think of philosophy as being made up of philosophical positions, philosophical questions and answers, or philosophical problems and solutions. My focus on philosophical accounts and theories does not exclude this. For the purposes of this paper, I use "philosophical account" and "philosophical theory" interchangeably and in a broad sense—though I am aware of the fact that some philosophers of science use "philosophical theory" in a stricter sense, for instance, as referring to sets of claims that are or consist of definitions that specify necessary and sufficient conditions.

A philosophical account or theory belongs to the philosophy of science if it makes claims about science. In other words, the object of philosophical theorizing in the philosophy of science is either some feature of science, such as systematicity, or a certain element of science, such as computer simulations, theoretical terms, model organisms, interventions, or causal inferences. We can say that a philosophical theory T concerns or is about some feature or element of science E. Figure 1 illustrates this way

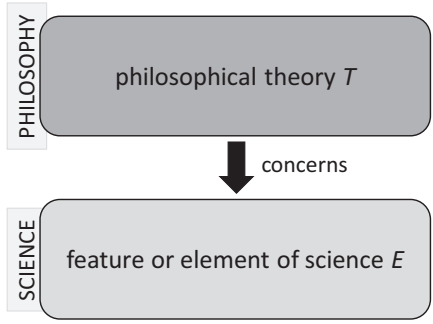


FIGURE 1. Philosophical theories in the philosophy of science

of conceptualizing the relation between philosophy of science and science.[2]

The expression "feature or element of science" is supposed to capture any entity that philosophers of science reason about. "Element of science" refers not only to material objects (for example, microscopes, electrons, scientists) but also to, for instance, scientific activities (such as modeling) and linguistic entities (such as observation statements). I speak of features and elements *of science*, not of scientific practice, because philosophers of science also make claims about hypothetical science, while "scientific practice" has the connotation of referring to actual science only (more on the notion of scientific practice in section 4). One might object that philosophers of science do not always make claims about science. Sometimes they are interested in metaphysical issues, which they take to be claims about the natural world itself rather than claims about how scientists investigate the natural world. I agree that we should not exclude metaphysical claims from the philosophy of science—especially if metaphysical questions are to be addressed in a naturalistic fashion (that is, by analyzing scientific knowledge). For reasons of simplicity, however, I will stick to the phrase "feature or element of science" and not always mention that philosophical theories might also concern the natural world studied by science.

3. Metanormativity

This section explicates the first of three ways in which philosophy of science can be normative. I argue that this kind of normativity, which I call "metanormativity," arises from the kind of claims that a philosophical

[2] Figure 1 is idealized in several respects. What is most important, it suggests that there exists a sharp boundary between philosophy (of science) and science, which is very difficult, if not impossible, to draw in reality.

theory contains. It is this kind of normativity that explains the manner in which the practice turn in the philosophy of science can be associated with a turn away from normative perspectives on science.

Let us start with an episode in the history of philosophy of science. The debate about epistemic reduction in the biological sciences started with the attempt to apply Nagel's (1961) formal model of theory reduction to biological cases, such as the putative reduction of Mendelian genetics to molecular biology (e.g., Hull 1974; Rosenberg 1985). It quickly became clear that Nagel's model encounters serious obstacles when applied to the biological sciences. As a response, Schaffner developed Nagel's model further and proposed his General Reduction-Replacement (GRR) model (1974, 1993). Schaffner explicitly constructs the GRR model as an ideal that need not be realized in contemporary scientific practice to be correct. He admits that the GRR model is only "peripheral" (1974, 111; 1993, 509) to biological practice because molecular biologists are not interested in obtaining the "complete chemical characterizations" (1974, 127) that are required for the kind of theory reductions he envisions. He treats the GRR model as a regulative ideal that should, but does not in fact, guide the development of molecular biology (1993, 511).[3] Other philosophers of science took the obstacles to applying Nagel's model to biology as evidence for antireductionism (Waters 1990). Nevertheless, almost all philosophers of biology agreed that Nagel's model was an adequate view of epistemic reduction in biology; at the time, it sounded "suspicious to change the standards of reduction" (Rosenberg 1985, 110). The situation changed in the 1990s when more and more philosophers realized that it does not make sense to impose an ill-fitting ideal of reduction on the biological sciences. Since then, several philosophers of biology have developed alternative accounts of epistemic reduction, which are based on extensive analyses of cases of epistemic reduction that actually occur in biological practice (for example, reductionist heuristics [Wimsatt 2006, 2007; Waters 2008], and reductive explanations [Sarkar 1998; Hüttemann and Love 2011; Kaiser 2015]). The debate about epistemic reduction in biology is only one example where the turn towards the empirical reality of scientific practice was accompanied by a turn away from philosophical accounts that construct ex cathedra normative claims about how science should be pursued or what it ideally looks like.

[3] In his recent work, Schaffner concedes that "what have traditionally been seen as robust reductions of one theory or one branch of science by another more fundamental one are largely a myth" (2006, 378). At first sight, this seems as an immense departure from his original position. Under closer inspection, however, one notices that Schaffner still regards the GRR model as an "ideal" (2006, 384) of what a complete reduction in biology would look like. For instance, he argues that in biology reductive, causal mechanical explanations are mere "*partial* reductions" and "reductions of the *creeping* sort" (2006, 397; emphasis in the original). This argument presupposes that there is an ideal of a complete, fully satisfying reduction.

3.1. Normative Claims About Science

What makes philosophical accounts, such as Schaffner's GRR model, normative? The object of Schaffner's account is epistemic reduction in biology (and medicine). This is the element of science *E* that the GRR model is concerned with. Schaffner does not describe what epistemic reduction in biology in fact is. Rather, he expresses how philosophers should understand the concept of epistemic reduction, states what good cases of epistemic reduction in biology are, and argues that biologists should try to achieve epistemic reductions that satisfy the requirements specified in the GRR model. His model is normative because it contains claims about its object of study (that is, epistemic reduction in biology), which are normative (rather than factual). We can generalize this thesis so that it holds for all philosophical accounts in the philosophy of science.

> *Metanormativity.* A philosophical theory *T* about a feature or element of science *E* is metanormative iff *T* contains normative claims about *E*.

This is the first way in which philosophy of science can be normative. I call this kind of normativity "metanormativity" because it arises from a feature of the philosophical account itself, rather than from how the account is developed or from what it is about. Figure 2 illustrates this kind of normativity.

What does it mean for a philosophical theory to contain normative claims about a feature or element of science? In general, normative claims can be evaluative statements and express the fact that something has or lacks a certain value, that something is good or bad, correct or incorrect. Normative claims can also be prescriptive and offer advice about what ought or ought not to be the case (for the distinction between evaluative and prescriptive norms see, e.g., McHugh 2012). In the philosophy of

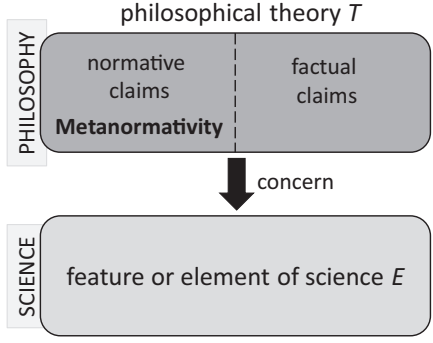


FIGURE 2. Metanormativity

science, normative claims about a certain feature or element of science are often normative in both the evaluative and the prescriptive sense. Schaffner's GRR model, for example, is metanormative in both senses. On the one hand, Schaffner's model contains the evaluative normative claim that only epistemic reductions that satisfy the GRR model are good or correct, other epistemic reductions being of the "*creeping* sort" (2006, 397; emphasis in the original) and thus lacking value. On the other, the GRR model is prescriptively metanormative because Schaffner assumes that philosophers should conceive of epistemic reduction according to his GRR model and that the model should guide biological research. Other examples of metanormative philosophical theories include Popper's (1959) view that falsifiability is the benchmark of science, Hempel's theory of what scientific explanation should look like (which is "not meant to describe how working scientists actually formulate their explanatory accounts" [1965, 412]), and Brandon's "normative ideal" (1996, 197) for adaptation explanations in evolutionary biology.

Normative claims are usually contrasted with factual (or positive or descriptive) claims that attempt to describe reality and thus are truth apt. Examples of philosophical theories that consist of only factual claims about their objects of study are Waters's (2008) analysis of investigative strategies in genetics and Winther's (2011) account of the integration of different kinds of part-whole explanations. These philosophical accounts are not metanormative, because they contain only factual claims about the element of science that they analyze. Waters describes how the central investigative strategy of classical genetics, the genetic approach, in fact works. Winther explicates how scientists actually develop and integrate different kinds of part-whole explanations of the tetrapod limb. Both describe science as it actually is, instead of making claims about science as it should be or about what is good science.

To conclude, one way in which philosophy of science can be normative is that it can make normative claims (that is, express evaluations and offer advice) about the feature or element of science that it studies. I call this kind of normativity "metanormativity" because it arises from a feature of the philosophical theory itself, namely, from the kind of claims that the philosophical theory contains.

3.2. Metanormativity After the Practice Turn

Having clarified the concept of metanormativity, I can now examine whether the practice turn in the philosophy of science is a turn away from metanormativity. If philosophical theories about science take into account the empirical reality of scientific practice, does that imply that they are not and cannot be metanormative? In what follows, I argue that the practice turn involves only a turn away from a specific style of metanormativity, which I call "ex cathedra metanormativity," but that

philosophy of science after the practice turn is compatible with other forms of metanormativity.

Philosophers of science who call for abandoning normative perspectives on science (e.g., Soler et al. 2014, 15) typically direct their criticism towards a very specific kind of metanormativity. They criticize philosophers who adopt a privileged viewpoint outside science and tell the practicing scientists what good science really is and how science is properly done. The objection is that these philosophers act as if they had "epistemic sovereignty" (Rouse 2002, 180), were "philosopher kings or philosophical police" (Sober 2008, xv), and were allowed to make metanormative claims about what is good science and how science should work ex cathedra (or from the ivory tower) without paying attention to how science in fact works. It should be emphasized that this is a caricature which very few, if any, philosophers of science fulfill. In a less radical version, metanormativity is ex cathedra if metanormative claims about what is good science and about how science ought to be pursued are developed and justified without taking into account actual scientific practice (this is what McMullin characterizes as "external philosophy of science" [1970, 24]). In other words, ex cathedra metanormative claims about a certain feature or element of science are not informed by and cannot fail in light of the empirical reality of scientific practice.

Examples of ex cathedra metanormativity can be found, for instance, in the metaphysics of science literature. Best-systems accounts of laws of nature (e.g., Lewis 1999, chaps. 1 and 15) state that generalizations in science should be regarded as laws only if they appear as axioms or theorems in the best (that is, the simplest and strongest) deductive system that contains everything we know in terms of natural properties. The metanormative claims in best-systems accounts are ex cathedra because they are said to be adequate independently of whether they capture actual cases of laws from scientific practice, whether they make sense of how scientists use the term "law," and whether all scientific knowledge can in fact be appropriately organized in the form of deductive systems and natural properties. That is, best-systems accounts of laws of nature exhibit ex cathedra metanormativity because they are not informed by and cannot fail in light of the empirical reality of scientific practice.

Ex cathedra metanormativity is thus not compatible with philosophers turning their attention to scientific practice, because "ex cathedra" means exactly the opposite, namely, ignoring scientific practice in developing and justifying a philosophical account. In contrast, metanormativity that is *not* in the ex cathedra style is perfectly compatible with the practice turn. Some philosophers even claim that metanormativity is an indispensable feature of *any* philosophical theory about science (e.g., Wimsatt 2007, 26; Sober 2008, xv).

What does it mean for philosophy of science to evaluate science and offer advice that is "contextual and sensitive to feedback" (Wimsatt 2007,

27) and that is informed by actual scientific practice? In my terminological framework, this means that a philosophical theory about a certain feature or element of science *E* contains only such normative claims about *E* that take into account, are drawn from, or are informed by factual claims about *E*. For example, Craver states that his descriptive project of characterizing mechanistic explanations in neurosciences "is the first step in a normative project: to clarify the distinction between good explanations and bad" (2007, viii). Similarly, Woodward emphasizes that his theory of causality and causal explanation also "makes recommendations about what one ought to mean by various causal and explanatory claims, rather than just attempting to describe how we use those claims" (2003, 7).

An interesting issue that I can only touch on here concerns the relation between normative and factual claims. Philosophers who aim at offering advice and evaluating science while paying close attention to the empirical reality of scientific practice face a challenge. On the one hand, they can only avoid an ex cathedra stance if they keep the relation between normative and factual claims about science as close as possible. Metanormative claims about some feature or element of science *E* should not be developed and justified independently from factual claims about *E*. Instead, metanormative claims about *E* should be based on or informed by factual claims about *E*. Simply deriving normative claims from factual claims, however, is illegitimate because it amounts to an is-ought fallacy (Bechtel and Richardson 2010, 10). On the other hand, philosophers of science might want to avoid the is-ought fallacy by developing and justifying their metanormative claims completely independently from factual claims about scientific practice (for example, by adducing a priori reasons [Schindler 2013]). If philosophers of science do this, however, their advice and evaluations become detached from the empirical reality of scientific practice, and the bugaboo of ex cathedra metanormativity looms again. To conclude, philosophers of science who have undergone the practice turn and still make metanormative claims need to meet this challenge and find ways to link their metanormative claims closely—but not too closely—to their factual claims. Promising approaches make use of, for example, the idea of a reflective equilibrium to specify how normative conclusions can be drawn from descriptive matters (Thagard 1988, chap. 7; cf. van Thiel and van Delden 2010).

4. Methodological Normativity

In this section, I analyze the methodology of philosophy of science in practice (PSP). My central claim is that even a philosophy of science that seeks to understand and accurately describe a certain element of scientific practice, and that thus contains only factual claims about science, is thoroughly normative. I call this kind of normativity "methodological

normativity" because it arises from the fact that the methodology of PSP inevitably involves normative assumptions.

4.1. The Methodology of Philosophy of Science in Practice

So far in this paper, we have been given only a rough idea of what it means for philosophy of science to pay close attention to actual scientific practice. Proponents of PSP have in common that they seek to understand science "from the inside while retaining a philosophical perspective" (Wimsatt 2007, 27), and that their philosophical accounts arise from an "empirical engagement with science" (Boumans and Leonelli 2013, 260). What do statements like these imply for the methodology of PSP?

Philosophy of science is a second-order discipline that studies the sciences that, in turn, study the natural world (e.g., McMullin 1970, 27; Sober 2008, xv).[4] For example, medical scientists aim at discovering the causes of complex diseases, such as cancer. Philosophers of science, in contrast, seek to understand, for instance, causal reasoning in cancer science as well as the strategies that cancer scientists employ to deal with causal complexity. In order to understand the methodology of PSP it is helpful to see in how far it presupposes a minimal methodological naturalism (cf. Giere 1999, 53–54; Bechtel 2008, 4–10). In my view, the methodology of PSP is similar to scientific methodology in at least one minor respect: in both fields, the theory or account that is developed must be empirically adequate, that is, it must capture and find evidential support in the available empirical data. This is why PSP is characterized as "empirical philosophy of science" (Mansnerus and Wagenknecht 2015, 38). Among the differences between the natural sciences and PSP is that empirical data in the former are about the natural world, whereas empirical data or information in the latter is about the natural sciences. That is, the empirical information against which a philosophical theory is "tested" is information from and about scientific practice. Scientists, in turn, develop scientific theories that they test against empirical data about the natural world (cf. Paul 2012).[5]

The claim that philosophy of science after the practice turn involves an empirical engagement with science can thus be specified as follows. In PSP, empirical information from and about scientific practice plays a central role in developing and justifying any philosophical account or theory. For example, when developing a philosophical account of causal inference

[4] This holds even for scientific metaphysics (e.g., Ross, Ladyman, and Kinkaid 2013), which studies scientific knowledge (for example, scientific theories or successful scientific practices) to draw metaphysical inferences.

[5] It is compatible with minimal methodological naturalism that the methodology of PSP and the methodology of the natural sciences differ in other respects. For example, one might claim that PSP is not a "science of science," because it is hermeneutic and proceeds through "acts of interpretation" (Schickore 2011, 461).

in oncology philosophers must take into account empirical information about the interventionist studies that scientists perform to identify carcinogens, about the causal explanations that oncologists develop, about the strategies that they employ to distinguish mere correlations from causal relations, and the new insights into the progression of cancer that are gained in these studies.

Empirical information that can be relevant to philosophical analysis is of diverse kinds. Accordingly, I understand the notion of scientific practice in a broad sense. It encompasses elements that are material or practical in a stricter sense (for example, scientific instruments, experiments, lab conditions, methods, and model organisms) as well as elements that are more theoretical, such as a scientist's epistemic activities (for example, explaining, testing, observing, modeling, theorizing, idealizing) and the results of these activities (for example, explanations, models, theories, generalizations).[6] Empirical information from and about scientific practice can also be information about the history of scientific practices (for example, in the form of historical case studies), which can lead to an integrated history and philosophy of science (cf. McMullin 1970; Schickore 2011; Kinzel 2015).

What does it mean for empirical information to play a central role in developing and justifying a philosophical theory? Philosophical accounts that pay close attention to scientific practice consist of factual claims that describe how science in fact works (in addition, they may contain metanormative claims as well; recall section 3). Unlike what the word "describes" suggests, philosophical accounts cannot be pure descriptions or one-to-one mappings of scientific practice. A philosophical account that is coherent and provides clarity and understanding cannot simply be read off scientific practice. Rather, it must result from a critical reconstruction of relevant empirical information from scientific practice. This is what Wimsatt seems to have in mind when he emphasizes that philosophical accounts must be developed from the inside of science while retaining a philosophical perspective (2007, 27), and this is what distinguishes PSP, for instance, from science journalism. In my view, the process of critically reconstructing relevant empirical information involves four major tasks: first, *selecting* empirical information from and about scientific practice that is relevant; second, *interpreting* empirical information, for example, by abstracting from irrelevant details and making explicit underlying assumptions; third, critically *evaluating* empirical information with the aim of establishing coherence; and fourth, *mutually adjusting* philosophical claims and empirical information until a reflective equilibrium is reached.

[6] Some authors put forward a narrower notion of scientific practice that includes material aspects of science only (Soler et al. 2014, 18) or that focuses on investigative practices (Waters 2014). In my view, theories and concepts remain important elements of scientific practice, which is why we should not exclude them.

I examine these four tasks in turn before addressing the question of which kind of normativity is involved in the process of critically reconstructing relevant empirical information from and about scientific practice.

First, the amount of empirical information that is available about most elements of scientific practice is enormous. Because of limited resources philosophers are forced to restrict their analyses to the empirical information that they think is particularly relevant to the question they address. The reasons why philosophers might pick out empirical information as relevant vary, but there are some general principles that guide the selection process. Typically, philosophers regard examples as relevant because they are paradigmatic or because they are representative for other cases of the same kind. By focusing on representative cases, philosophers try to make sure that the philosophical theory they develop holds not only for the analyzed cases but also for science or a scientific field in general. Another major reason for assessing a case as relevant is that it is of particular importance to a scientific field, for instance, if it has driven scientific research for a longer period of time or if it is an example for how the success (or failure) of research in that field is promoted. Since success is seen as a central goal of science, it is also of particular interest to philosophers (Giere 1999, 53; Norton 2003, 648). Successful examples might, for instance, be those that appear in established textbooks or that are much discussed in a certain field. A case might also be assessed as relevant because it contributes to achieving another central goal of a scientific field, such as manipulation or disease control, or because it concerns central processes of life or of our world (for example, reproduction of living beings, quantum entanglement, and the Big Bang).

Second, the process of developing a philosophical theory while taking relevant empirical information into account is often not straightforward but involves a great deal of abstraction, explication, and "interpretation" (Schickore 2011, 471). Philosophers must explicate background assumptions that scientists implicitly presuppose in their experimentation and reasoning, they must establish connections between seemingly unrelated claims and concepts, they must abstract from philosophically irrelevant details, and they must draw philosophical inferences from empirical information. Consider the example of developing an account of what makes biological explanations reductive (Hüttemann and Love 2011; Kaiser 2015). The first challenge that philosophers encounter is that only very few biologists indicate whether the explanations they give are reductive or not. They argue about the adequacy of explanations but not about their reductive character because this is just not important to them. Some biologists engage in intensive debates about the "limits of reductionism" (Mazzocchi 2008, 10) and the need to move "beyond reductionism" (Gallagher and Appenzeller 1999, 79). But even then biologists rarely speak about reductive explanations. Rather, they discuss the correctness of a reductionist approach and the adequacy of applying reductive methods. Sometimes

biologists do not even use terms containing "reduc-" but, nevertheless, express assumptions about reductive explanation, such as when they discuss part-whole explanations or the method of decomposition. This example illustrates the extent to which the process of developing philosophical claims on the basis of empirical information requires acts of abstraction, explication, and interpretation.

Third, philosophers are confronted with many differences or even inconsistencies within and among scientific fields. For instance, explanatory and investigative strategies vary, the same concepts are understood differently, and different background assumptions are made. To develop a coherent, unified theory about a certain feature or element of scientific practice, philosophers must take up a critical stance and sort out the empirical information that can be dismissed as false, misleading, or biased. For example, what biologists mean by "reductionism" and what they think constrains the adequacy of a reductive explanation is by no means homogenous and involves inconsistencies. Some biologists identify reductive explanation with additive explanations, that is, with explanations in which biological systems are treated as aggregative systems (Kitano 2002, 1662). Other biologists explicitly reject this claim because it results in a too restricted view of reductive explanation. They state that "[m]olecular biologists . . . do not hold the naive view that complex structures and processes are just sums of their parts" (Fincham 2000, 343). If one wants to develop a coherent theory of reductive explanation, one needs to ponder which of these claims should inform the philosophical theory (for example, because they are more common or are best in line with other relevant empirical information) and which should be sorted out as incorrect, rare, too vague, or insufficiently justified.

Finally, as the other three tasks already indicate, the process of developing a philosophical theory by taking into account empirical information from and about scientific practice is not a one-way process but involves a repeated mutual adjustment and moving back and forth between philosophical theory and empirical information (this is why figure 3 includes arrows leading from science via empirical information to philosophy and back again). This process can also be characterized as an inherently hermeneutic endeavor (Schickore 2011) and as an iterative (inner) dialogue between abstract theory and concrete data (Mansnerus and Wagenknecht 2015). The process of mutual adjustment often starts with provisional philosophical claims and preconceptions that are brought together with, sharpened, and modified in the light of provisional selections, interpretations, and evaluations of empirical information. The process comes to an end, for example, as soon as a reflective equilibrium between philosophical theory and empirical information is reached (Thagard 1988, 119; for the general idea see, e.g., Elgin 1996, chap. 4).

To conclude, in the philosophy of science after the practice turn, developing a philosophical theory *T* about some feature or element of science *E*

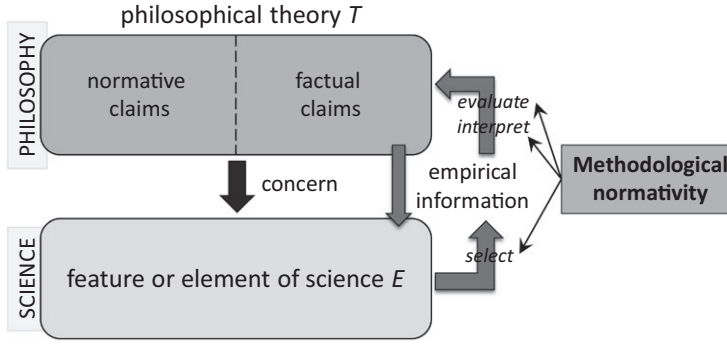


FIGURE 3. Methodological normativity

involves selecting, interpreting, and evaluating relevant empirical information about *E* from and about scientific practice as well as making mutual adjustments between philosophical theory and empirical information.

4.2. Methodological Normativity

The goal of this section is to show that even empirically based philosophy of science, which seeks to understand and describe how scientific practice in fact works, is thoroughly normative. The kind of normativity involved here is different from metanormativity because the normative claims do not concern the object of philosophical theorizing itself. Rather, they concern the methodology by which an empirically based philosophical theory is developed. This is why I refer to this kind of normativity as "methodological normativity." Figure 3 illustrates this kind of normativity.

If philosophers develop a philosophical theory by selecting, interpreting, and evaluating relevant empirical information and mutually adjusting philosophical claims and empirical information, they presuppose, usually implicitly, certain methodological norms. These are norms that, for instance, express what is good empirical information (relative to the philosophical question at stake) and that give advice about how philosophers should proceed in developing a particular theory through selecting, interpreting, and evaluating empirical information and mutually adjusting theory and empirical information. The following examples of methodologically normative assumptions illustrate what these methodological norms are that guide theory development in PSP:

> An example of *E* is good because it is an instance of successful science/contributes to a major aim of a scientific field (for example, manipulation, disease control, prediction, or technological progress).

An example of *E* is good because it is paradigmatic/clear-cut/robust/representative for cases of the same kind.
Empirical information about *E* from a certain scientific field is good because this field is especially successful.
Empirical information about another element *E** should be taken into account by a philosophical theory *T* about *E* because *E* and *E** are closely related.
Claims of scientists about *E* should be excluded from the empirical basis of *T* because these claims are incorrect/biased/too vague.
Empirical information about *E* should not be taken into account by *T* because it would prevent establishing coherence.
A claim of philosophical theory *T* should be revised/abandoned because it conflicts with relevant empirical information about *E*.

As presented above, some of these methodological norms are formulated as evaluative normative claims, others as prescriptive. Nevertheless, all claims seem easily translatable from one formulation to another. This reflects the fact that methodological norms that figure in philosophy of science after the practice turn are typically both evaluative and prescriptive. That is, they express that certain kinds of empirical information or philosophical procedures have or lack a value, and they offer advice about what philosophers ought or ought not to do (for example, which empirical information they should take into account and which they should ignore).

In most cases, methodologically normative assumptions will not be explicitly stated but rather will implicitly guide the selection, interpretation, and evaluation of empirical information and the process of mutually adjusting philosophical theory and empirical information. Furthermore, we need not assume that these methodological norms are static. They can change over time, if we, for instance, learn how to better assess or select empirical information from scientific practice.

To conclude, philosophical theories that make factual claims about a certain feature or element of science may not be metanormative (if they contain factual claims only; recall section 3). Still, these theories are normative because the philosophical methodology of developing factual claims about science while taking into account empirical information from and about scientific practice inevitably involves making (implicit) normative assumptions about how to select, interpret, and evaluate empirical information and how to mutually adjust theory and empirical information. I call this second kind of normativity "methodological normativity" because it arises from the methodology by which an empirically based philosophical theory about science is developed.

Methodological normativity. A philosophical theory *T* about a feature or element of science *E* is methodologically normative iff *T* contains factual claims

about *E* that (implicitly) rely on normative claims about how to select, interpret, evaluate, and mutually adjust empirical information about *E* in developing *T*.

Methodological normativity differs from metanormativity in two important ways. First, methodological normativity stems from norms that concern the philosophical methodology and that thus commit philosophers, not scientists, to handling empirical information in a certain way or to seeking coherence between philosophical theory and empirical information in a certain way. Metanormativity, in contrast, commits primarily scientists to, for instance, seeking theory reductions à la Schaffner or adopting a specific understanding of the concept of a mechanism. Second, methodological normativity does not require that a philosophical theory contains normative claims about its object of study *E* (as metanormativity does). Rather, methodological normativity applies to empirically based philosophy of science only, that is, to philosophical theories that contain factual claims about *E*. If a philosophical theory contains factual and normative claims about *E* it is methodologically normative and metanormative (see section 3.2). In sum, introducing the category of methodological normativity reveals in what way even a philosophical theory that describes how science in fact works is thoroughly normative.

5. Object Normativity

In this section, I identify a third way in which philosophy of science can be normative: object normativity. This kind of normativity emerges from the fact that the object of philosophical theorizing itself can be normative. This is the case if philosophers reason about epistemic or social norms and their roles in science.[7] Among the questions that are of philosophical interest is, for instance, the question of whether epistemic norms such as simplicity, precision, explanatory power, and predictive success guide how scientists identify their objects of study, interpret empirical data, and choose between competing theories or explanations (e.g., Kuhn 1962). Philosophers of science also controversially discuss whether social or political norms, such as democracy, human rights, or gender biases, influence the scientific process of acquiring knowledge about the natural world and may jeopardize the objectivity of scientific knowledge (Longino 1990; Kitcher 2011). Because this kind of normativity arises from the objects of a philosophical theory being norms (or being related to norms) I refer to it as "object normativity."

[7] I speak about norms in science, rather than about values, because the concept of a norm is broader and accounts for evaluative as well as for prescriptive normative claims in science. I understand "social norms" in a broad way including various kinds of non-epistemic norms that are relevant to society.

Object normativity. A philosophical theory *T* about a feature or element of science *X* is object normative iff *T* refers to epistemic or social norms in science.

In the following two subsections, I explicate what it can mean for a philosophical theory to refer to epistemic or social norms in science and thereby introduce different subtypes of object normativity. First, I distinguish philosophical theories that describe which norms are in fact accepted in scientific practice from philosophical theories that posit norms that should be accepted in science (section 5.1). Second, among philosophical theories that refer to epistemic norms in science, I distinguish those that concern epistemic norms themselves from philosophical theories that are about some nonnormative element of science but that relate this element to certain epistemic norms (section 5.2).

Before moving on let me add a final general remark. One might wonder how object normativity relates to the turn towards philosophical accounts that engage with actual scientific practice. The practice turn is said to involve a shift to perspectives on science that are "more realistic" (Soler et al. 2014, 18), for instance, because they recognize the deep intertwinement of science and society. In line with this, Rouse warns against construing science as "clearly bounded and distinct from extrascientific 'context'" (2002, 164). Philosophers of science in practice typically avoid this danger because they take into account the practice of scientific research in its entire variety, including processes of inquiry, institutional settings, and social dynamics among investigators (Boumans and Leonelli 2013). This includes, for example, recognizing the political dimension of knowledge and the ways in which scientific fields may be shaped by the uses to which scientific knowledge may be put, such as how gene patents affect medical genetic testing (Carrier, Howard, and Kourany 2008).

Accordingly, some philosophers of science in practice argue that examining the goals of scientific activities requires not only epistemological considerations but also reflections on "the values, norms, and ideals inherent in the pursuit of scientific knowledge" (Ankeny et al. 2011, 305). This might suggest that any philosophical theory in the philosophy of science in practice *must* refer to epistemic or social norms and thus be object normative. I think that this claim is too strong and that we gain nothing from imposing such a strict requirement on what is seen as "proper" practice-oriented philosophy of science. Nevertheless, from a general perspective, the practice turn is accompanied by a shift to a philosophy of science that is more object normative.

5.1. Describing Norms Versus Positing Norms

The first distinction of subtypes of object normativity results from linking object normativity to metanormativity. In Section 3, I distinguished philosophical theories that contain normative claims about their objects

of study *E* (and thus are metanormative) from those that contain only factual claims about *E*. This difference applies to any feature or element of science—to nonnormative elements of science, such as causal inferences, model organisms, and mathematical equations, as well as to epistemic and social norms that figure in science (which give rise to object normativity), such as explanatory power or gender biases. Figure 4 illustrates the four possible types of philosophical theories that result from combining the two different kinds of claims that a philosophical theory can contain with the two kinds of objects that the theory can be concerned with.

Combinations ① and ② represent philosophical theories that are not object normative, because they refer to nonnormative elements of science only. Philosophical theories of type ① contain normative claims about nonnormative elements of science (for example, the claim that only cases of reductions that fulfill Schaffner's GRR model are good) and thus are metanormative. By contrast, philosophical theories of type ② make only factual claims about nonnormative elements of science and thus are neither metanormative nor object normative. The difference between ① and ② was spelled out in detail in section 3.

Consider now philosophical theories that refer to epistemic or social norms in science and thus are object normative (combinations ③ and ④). Combination ④ represents philosophical theories that describe which epistemic or social norms are in fact accepted in scientific practice. These philosophical theories make factual claims about which norms actually influence scientific inquiry. I refer to this as "describing norms." Examples of descriptions of norms in science are Lloyd's (2006) theory of how gender biases influence the development of adaptive explanations of female orgasm and my analysis of how biologists evaluate reductive explanations

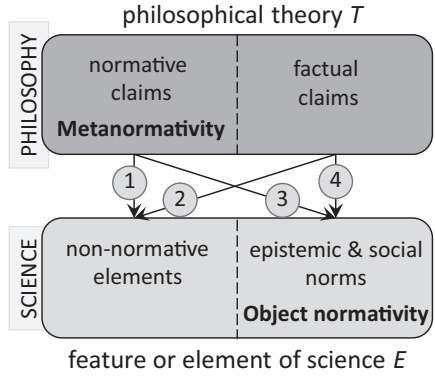


FIGURE 4. Possible combinations of metanormativity and object normativity

as adequate, that is, which norms of reductive explanation they accept (Kaiser 2015). By contrast, combination ③ represents philosophical theories that posit which epistemic or social norms should apply to science and justify why these norms should be accepted. I refer to this as "positing norms." Philosophical theories of this kind refer to norms by making normative claims about which epistemic or social norms ideally should influence, for instance, the process of gaining scientific knowledge (and how they should do so). Examples of normative claims about norms in science include: Kitcher's theory of well-ordered science, which constructs an idealistic picture of how decisions about the significance of research projects should be democratically assessed (2011); philosophical theories of simplicity as the best criterion for choosing among competing scientific theories and explanations (Sober 1975; Thagard 1988; White 2005); and Craver's (2007) account of how mechanistic explanations in neuroscience should be evaluated. In sum, philosophical theories that describe accepted norms (that is, ④) are as object normative as philosophical theories that posit which norms should be accepted (that is, ③), but only the latter are also metanormative because only they contain normative claims about the norms to which they refer.[8]

5.2. Theorizing About Norms Versus Relating to Norms

In this section, I reveal a distinction between two subtypes of object normativity that applies only to philosophical theories that refer to *epistemic* norms. This distinction emerges from the fact that claims about epistemic norms in science—whether factual or normative—can figure differently in a philosophical theory. On the one hand, epistemic norms can be the objects of philosophical theorizing. In these cases, the philosophical theory is *about* these norms. For example, Thagard (1988) has proposed a theory about the epistemic value of simplicity, considering how we should understand it and why it is justified. Similarly, Lloyd's (2006) analysis is about gender biases and how they affect the development of adaptive explanations of female orgasm. I refer to these cases as "theorizing about norms."

On the other hand, reference to epistemic norms can be less central to a philosophical theory. In these cases, the object *E* of a philosophical theory is not epistemic norms but another nonnormative element of science (for example, causal inference, reduction, or the concept of a gene), and the philosophical theory includes claims about how *E* is related to certain epistemic norms (that either are in fact or should be accepted in science).

[8] One might express the difference between ③ and ④ also by claiming that philosophical theories of type ④ refer to intrinsic norms (that is, norms that are inherent in scientific practice), whereas those of type ③ refer to extrinsic norms (that is, norms that are posited from outside science).

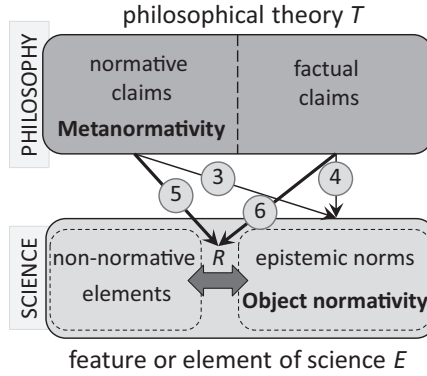


FIGURE 5. Subtypes of object normativity[9]

For example, philosophical analyses of the concept of reductive explanation are about reductive explanations but they may also elucidate why reductive explanations are adequate with respect to some phenomena and inadequate with respect to others, that is, how the reductive character of an explanation affects the epistemic value of explanatory success. This is an example of what I refer to as "relating to norms."[10] Figure 5 illustrates the difference between object normativity as theorizing about norms and as relating to norms.

Object-normative philosophical theories can either directly address epistemic norms in science, that is, theorize about norms (③ and ④), or they can address another element of science *E* and its relation *R* to certain epistemic norms, that is, they can relate *E* to epistemic norms (⑤ and ⑥). This is the difference between object normativity as theorizing about norms and as relating to norms. These two subtypes of object normativity can be combined with the other two subtypes of object normativity, which I introduced in section 5.1, namely, describing norms and positing norms. Object-normative philosophical theories can either make factual claims

[9] Philosophical theories that do not refer to epistemic or social norms (① and ② in figure 4) are not object normative and thus are omitted from figure 5.

[10] To which of these two categories a philosophical theory belongs depends on how fine-grainedly it is individuated. For instance, Craver's account of mechanistic explanation is not about epistemic norms. Still, his account includes claims about what the norms of mechanistic explanations are and should be (Craver 2007, 20 and 111). If we conceptualized these claims as a separate theory, it would be a case of theorizing about norms, otherwise Carver's account of mechanistic explanation would be a case of relating to norms.

about epistemic norms or about the relations to epistemic norms (④ and ⑥), or they can make normative claims about these norms or relations (③ and ⑤), which gives rise to metanormativity (see section 5.1).

Recognizing the difference between object normativity as theorizing about norms and as relating to norms enables us to assess whether and to what extent object normativity is or should be prevalent in practice-oriented philosophy of science. Some philosophers hold the apparently strong view that any philosophical theory about science should refer to epistemic norms and thus should be object normative. For instance, Waters argues that philosophers of science should provide an understanding of "how the sciences work (and don't work) with respect to epistemic virtues that we value" (2004, 48). A philosophical theory of what genes are, for example, should capture not only how scientists reason about genes, investigate genes, and use gene terminology. It also must clarify how the gene concept is related, for instance, to the epistemic norm of explanatory power (for example, by discussing the limitations of gene-based explanations) or to the epistemic norm of investigative utility (for example, by revealing the usefulness of chief methods in gene-centered sciences). At first sight, the claim that any philosophical theory about science must refer to epistemic norms seems to result in a too restrictive view of what proper philosophy of science is. If we apply the distinction between relating to norms versus theorizing about norms, however, we see that the claim is weaker and more plausible. Waters does not call for a philosophy of science that studies epistemic norms in science only (which would be object normative as theorizing about norms). Besides epistemic norms, there are plenty of other features and elements of science that are worthy of philosophical investigation. Waters's claim is that any philosophical theory about these other features and elements of science (such as gene-based explanations) must explicate how they relate to epistemic norms. Regardless of whether one thinks that Waters's claim that object normativity should be prevalent in the philosophy of science in practice is fully convincing or not, my analysis shows that this claim must be understood to concern a specific subtype of object normativity, namely, object normativity as relating to norms.

6. Interrelations Between the Three Kinds of Normativity

The goal of this section is to explicate how the three kinds of normativity that characterize the philosophy of science relate to each other and can be combined.

Depending on the kinds of claims that compose a philosophical theory and depending on the kinds of objects with which a philosophical theory is concerned, a philosophical theory falls into one of four groups. It can be (1) metanormative (if it makes normative claims about non-normative elements of science), (2) object normative (if it makes factual

claims about epistemic or social norms in science), (3) metanormative and object normative (if it makes normative claims about epistemic or social norms in science), or (4) neither metanormative nor object normative (if it makes only factual claims about nonnormative elements in science) (see section 5.1). All four combinations are possible because metanormativity and object normativity are not only distinct types of normativity but also independent of each other. The kind of claims that a philosophical theory makes does not commit it to a specific kind of object of philosophical theorizing (and vice versa).

Now, whenever a philosophical theory contains factual claims about some feature or element of science, it is methodologically normative because developing these factual claims on the basis of empirical information from and about scientific practice presupposes (implicit) normative claims about how to select, interpret, evaluate, and mutually adjust empirical information (see section 4.2). This holds for factual claims about nonnormative elements of science as well as for factual claims about norms in science. Hence, methodological normativity can but need not be combined with object normativity. It also holds for philosophical theories that contain not only factual claims but also normative claims (see section 3.2). The only kind of normativity that methodological normativity is incompatible with is metanormativity of the ex cathedra style. Methodological normativity requires that a philosophical theory is empirically based and makes factual claims about actual scientific practice, which ex cathedra metanormativity rejects. By contrast, a metanormative theory that is sensitive to the empirical reality of scientific practice relies on methodological normativity because its normative claims must be informed by or connected in some other way to factual claims (see section 3.2). It is thus possible that a philosophical theory possesses all three kinds of normativity: a philosophical theory that makes factual as well as normative claims about epistemic or social norms in science (either by making claims about norms or by relating some nonnormative element to epistemic norms; see section 5.2) is methodologically metanormative and object normative.

One might wonder whether methodological normativity involves or depends on claims that are, themselves, metanormative or object normative. Typical methodological norms that implicitly guide theory development in the philosophy of science in practice make use of, for instance, specific notions of scientific success, and they rely on assumptions about what the goals of science are or should be. Hence, one could argue that methodological normativity presupposes metanormativity and object normativity. I think this claim is basically correct. Two points must be emphasized, however: first, this claim holds only for some methodological norms; second, this claim does not imply that methodological normativity reduces to metanormativity or to object normativity and can be eliminated from my typology of normativity in the philosophy of science.

Methodological normativity arises from normative claims that tell philosophers which types of empirical information are valuable (such as paradigmatic examples, instances of successful science, unbiased and clear-cut empirical information) and how they should proceed in developing a philosophical theory (for example, establish coherence, reveal conceptual connections, ignore biased or nonrepresentative cases). Methodological normativity is thus distinct from metanormativity because methodological norms tell philosophers what to value and what to do, whereas metanormativity concerns norms that tell scientists what to value and what to do (for example, what good cases of reductions are or which methods to apply). Methodological normativity is also distinct from object normativity because object normativity arises from norms in science, whereas methodological normativity traces back to norms in philosophy.

Despite the distinctness of all three kinds of normativity, some methodological norms presuppose assumptions about which epistemic norms are or should be accepted in science. An example is the methodological norm that a case of *E* is good and should be taken into account by philosophers because it plays a crucial role in achieving the aim of manipulating a specific disease. This methodological norm rests on the assumption that manipulation is accepted as a major epistemic norm in medical science. Other methodological norms presuppose assumptions about what should be regarded as scientific success, such as the methodological norm that empirical information about *E* from a certain scientific field is good because this field is especially successful in terms of making novel predictions. This methodological norm depends on the metanormative claim that the notion of success should be spelled out in terms of novel predictions. Hence, some methodologically normative claims rely on object-normative or metanormative claims, even though they are not themselves object normative or metanormative.

7. Conclusion

The goal of this paper is to draw attention to the fact that normativity in philosophy of science is a multifaceted phenomenon. Normativity is not a single feature that a philosophical theory either has or doesn't have. The paper articulates three different ways in which a philosophical theory about science can be normative. Each of the three kinds of normativity has a different origin. Methodological normativity emerges from norms involved in the philosophical methodology of developing factual claims about scientific practice. Object normativity is due to the normativity of the objects of philosophical theorizing. Metanormativity arises from the normativity of the claims that a philosophical theory contains.

Even though the practice turn in the philosophy of science is sometimes characterized as a "shift from normative to descriptive perspectives

on science" (Soler et al. 2014, 15) my analysis shows that it is only a shift from metanormativity of the ex cathedra sort. The turn to scientific practice involves a move away from the approach of giving advice about what science really is and how to do proper science without considering the empirical reality of scientific practice. Other than ex cathedra styles, then, philosophy of science in practice is compatible with metanormativity. What is more, the practice turn is said to be a shift to more realistic views of science that also recognize the norms inherent in the pursuit of scientific knowledge and the deep intertwinement of science and society. Frequently, the practice turn will thus involve a turn to object normativity. Finally, even if philosophers of science in practice are not interested in making normative claims (metanormativity) or in reflecting on epistemic or social norms in science (object normativity), their theories about scientific practice will not be free from normativity, because their methodology inevitably involves normative assumptions about how to select, interpret, and evaluate empirical information, and how to mutually adjust philosophical theories and empirical information.

Bielefeld University
Department of Philosophy
Postfach, 100131
33501 Bielefeld
Germany
kaiser.m@uni-bielefeld.de

Acknowledgments

This project was partly funded by the German Research Foundation (FOR 1063). My special thanks go to Ken Waters, who discussed various versions of this paper with me. His ideas and suggestions always enhance my philosophical thinking. I would also like to thank my colleagues at the University of Cologne and at Bielefeld University for their helpful comments on previous versions of the paper.

References

Ankeny, R., H. Chang, M. Boumans, and M. Boon. 2011. "Introduction: Philosophy of Science in Practice." *European Journal for Philosophy of Science* 1:303–7.

Bechtel, W. 2008. "Mechanisms in Cognitive Psychology: What Are the Operations?" *Philosophy of Science* 75:983–94.

Bechtel, W., and R. C. Richardson. 2010. *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*. Cambridge, Mass.: MIT Press.

Boumans, M., and S. Leonelli. 2013. "Introduction: On the Philosophy of Science in Practice." *Journal for General Philosophy of Science* 44, no. 2:259–61.
Brandon, R. N. 1996. *Concepts and Methods in Evolutionary Biology*. Cambridge: Cambridge University Press.
Carrier, M., D. Howard, and J. Kourany, eds. 2008. *The Challenge of the Social and the Pressure of Practice: Science and Values Revisited*. Pittsburgh, Penn.: University of Pittsburgh Press.
Chang, H. 2011. "The Philosophical Grammar of Scientific Practice." *International Studies in the Philosophy of Science* 25, no. 3:205–21.
Craver, C. F. 2007. *Explaining the Brain: Mechanisms and the Mosaic Unity of Neuroscience*. Oxford: Oxford University Press.
Elgin, C. Z. 1996. *Considered Judgment*. Princeton: Princeton University Press.
Fincham, J. R. S. 2000. "Reductionism Should Be Clarified, Not Dismissed." *Nature* 406:343.
Gallagher, R., and T. Appenzeller. 1999. "Beyond Reductionism." *Science* 284:79.
Giere, R. N. 1999. *Science Without Laws*. Chicago: University of Chicago Press.
Hempel, C. G. 1965. *Aspects of Scientific Explanation: And Other Essays in the Philosophy of Science*. New York: Free Press.
Hull, D. 1974. *The Philosophy of Biological Science*. Upper Saddle River, N.J.: Prentice-Hall.
Hüttemann, A., and A. C. Love. 2011. "Aspects of Reductive Explanation in Biological Science: Intrinsicality, Fundamentality, and Temporality." *British Journal for Philosophy of Science* 62, no. 3:519–49.
Kaiser, M. I. 2015. *Reductive Explanation in the Biological Sciences*. Cham: Springer.
Kinzel, K. 2015. "Narrative and Evidence: How Can Case Studies from the History of Science Support Claims in the Philosophy of Science?" *Studies in History and Philosophy of Science* 49:48–57.
Kitano, H. 2002. "Systems Biology: A Brief Overview." *Science* 295:1662–64.
Kitcher, P. 2011. *Science in a Democratic Society*. Amherst, N.Y.: Prometheus Press.
Kuhn, T. 1962. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.
Lewis, D. 1999. *Papers in Metaphysics and Epistemology*. Cambridge: Cambridge University Press.
Lloyd, E. A. 2006. *The Case of the Female Orgasm: Bias in the Science of Evolution*. Cambridge, Mass.: Harvard University Press.
Longino, H. 1990. *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry*. Princeton: Princeton University Press.

Mansnerus, E., and S. Wagenknecht. 2015. "Feeling with the Organism: A Blueprint for an Empirical Philosophy of Science." In *Empirical Philosophy of Science: Introducing Qualitative Methods into Philosophy of Science*, edited by S. Wagenknecht, N. J. Nersessian, and H. Andersen, 37–61. Cham: Springer.

Mazzocchi, F. 2008. "Complexity in Biology: Exceeding the Limits of Reductionism and Determinism Using Complexity Theory." *EMBO Reports* 9:10–14.

McHugh, C. 2012. "The Truth Norm of Belief." *Pacific Philosophical Quarterly* 93:8–30.

McMullin, E. 1970. "The History and Philosophy of Science: A Taxonomy." In *Historical and Philosophical Perspectives on Science*, edited by R. H. Steuwer, 12–67. *Minnesota Studies in the Philosophy of Science 5*. Minneapolis, MN: University of Minnesota Press.

Nagel, E. 1961. *The Structure of Science: Problems in the Logic of Scientific Explanation*. London: Routledge.

Norton, J. 2003. "A Material Theory of Induction." *Philosophy of Science* 70, no. 4:647–70.

Paul, L. A. 2012. "Metaphysics as Modeling: The Handmaiden's Tale." *Philosophical Studies* 160:1–29.

Popper, K. 1959. *The Logic of Scientific Discovery*. Translation of Logik der Forschung. London: Hutchinson.

Rosenberg, A. 1985. *The Structure of Biological Science*. Cambridge: Cambridge University Press.

Ross, D., J. Ladyman, and H. Kincaid, eds. 2013. *Scientific Metaphysics*. Oxford: Oxford University Press.

Rouse, J. 2002. *How Scientific Practices Matters: Reclaiming Philosophical Naturalism*. Chicago: University of Chicago Press.

Sarkar, S. 1998. *Genetics and Reductionism*. Cambridge: Cambridge University Press.

Schaffner, K. F. 1974. "The Peripherality of Reductionism in the Development of Molecular Biology." *Journal of the History of Biology* 7, no. 1:111–39.

———. 1993. *Discovery and Explanation in Biology and Medicine*. Chicago: University of Chicago Press.

———. 2006. "Reduction: The Cheshire Cat Problem and a Return to the Roots." *Synthese* 151:377–402.

Schickore, J. 2011. "More Thoughts on HPS: Another 20 Years Later." *Perspectives on Science* 19, no. 4:453–81.

Schindler, S. 2013. "The Kuhnian Mode of HPS." *Synthese* 190, no. 18:4137–54.

Sober, E. 1975. *Simplicity*. Oxford: Clarendon Press.

———. 2008. *Evidence and Evolution: The Logic Behind the Science*. Cambridge: Cambridge University Press.

Soler, L., S. Zwart, M. Lynch, and V. Israel-Jost, eds. 2014. *Science After the Practice Turn in the Philosophy, History, and Social Studies of Science*. New York: Routledge.

Thagard, P. 1988. *Computational Philosophy of Science*. Cambridge, MA: MIT Press.

van Thiel, G., and J. van Delden. 2010. "Reflective Equilibrium as a Normative Empirical Model." *Ethical Perspectives* 17:183–202.

Waters, C. K. 1990. "Why the Anti-Reductionist Consensus Won't Survive: The Case of Classical Mendelian Genetics." *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 125–139. Chicago: University of Chicago Press.

———. 2004. "What Concept Analysis in Philosophy of Science Should Be (and Why Competing Philosophical Analyses of Gene Concepts Cannot Be Tested by Polling Scientists)." *History and Philosophy of the Life Sciences* 26, no. 1:29–58.

———. 2008. "Beyond Theoretical Reduction and Layer-Cake Antireduction: How DNA Retooled Genetics and Transformed Biological Practice." In *The Oxford Handbook of Philosophy of Biology*, edited by M. Ruse, 238–62. Oxford: Oxford University Press.

———. 2014. "Shifting Attention from Theory to Practice in Philosophy of Biology." In *New Directions in the Philosophy of Science, vol. 5 of The Philosophy of Science in a European Perspective*, edited by M. C. Galavotti, D. Dieks, W. J. Gonzalez, S. Hartmann, T. Uebel, and M. Weber, 121–39. Cham: Springer.

White, R. 2005. "Why Favour Simplicity?" *Analysis*, 65, 205–10.

Wimsatt, W. C. 2006. "Reductionism and Its Heuristics: Making Methodological Reductionism Honest." *Synthese*, 151, no. 3:445–75.

———. 2007. *Re-Engineering Philosophy for Limited Beings: Piecewise Approximations to Reality*. Cambridge, Mass.: Harvard University Press.

Winther, R. G. 2011. "Part-Whole Science." *Synthese* 178:397–427.

Woodward, J. 2003. *Making Things Happen: A Theory of Causal Explanation*. Oxford: Oxford University Press.

Disability & Society

ISSN: 0968-7599 (Print) 1360-0508 (Online) Journal homepage: https://www.tandfonline.com/loi/cdso20

Exploring normativity in disability studies

Simo Vehmas & Nick Watson

To cite this article: Simo Vehmas & Nick Watson (2016) Exploring normativity in disability studies, Disability & Society, 31:1, 1-16, DOI: 10.1080/09687599.2015.1120657

To link to this article: https://doi.org/10.1080/09687599.2015.1120657

Published online: 08 Jan 2016.

DISABILITY & SOCIETY, 2016
VOL. 31, NO. 1, 1–16
http://dx.doi.org/10.1080/09687599.2015.1120657

Routledge
Taylor & Francis Group

Exploring normativity in disability studies

Simo Vehmas[a] and Nick Watson[b]

[a]Disability Studies, University of Helsinki, Helsinki, Finland; [b]Disability Studies, University of Glasgow, Glasgow, UK

ABSTRACT

Normativity is a concept that is often misapplied in disability studies, especially in 'postconventional' accounts, where the concept is conflated with 'normal', 'normate', or 'standard'. This article addresses this confusion, explores the meaning and use of 'normativity', and presents some analytic tools to discuss normative issues of right and wrong. The article finishes by discussing examples where conceptual confusions result in confused normative judgments focusing in particular on agency, responsibility and moral status. The article argues that disability research should carefully consider the use of theories and empirical knowledge in the light of their ethical implications as well as the lived experiences of disability.

ARTICLE HISTORY
Received 4 May 2015
Accepted 12 November 2015

KEYWORDS
Ethics; moral responsibility; moral status; posthumanism

Points of interest

- This article explores the uses of normativity in disability studies.
- It is argued that theorists in postconventional disability studies have misconstrued the meaning of normativity and incorrectly conflated it with normality.
- The article argues for the importance of normative, especially ethical, engagement in relation to disability, and provides some conceptual tools for the examination of moral right and wrong.
- Posthumanist appeals for collective responsibility, and especially for the pursuit to nullify the separation between humans and animals, may put many disabled people at risk.
- Finally, the article offers insights about the ethical implications of the use of theories and empirical claims.

Introduction

In their recent paper critiquing critical disability studies, Vehmas and Watson (2014) expressed concern about the theoretical accuracy and practical feasibility of this approach. One area they specifically concentrated on was the neglect of the normative; namely, the ethical and political dimensions that relate to disability. They argued that an adequate examination of

CONTACT Simo Vehmas simo.vehmas@helsinki.fi

disability requires engagement with such moral and political issues. This is because disability both as a phenomenon and as a concept is in its essence normative; it expresses normative ideas and assumptions concerning what kinds of capacities or possibilities people should have or be afforded in order to lead a good life, and/or how society ought to be organised in order to treat its members equally and fairly.

In the traditional medical, individualistic approach, disability is the result from impaired functioning or being and is ameliorated through the promotion of interventions that remove or reduce the impairment. In contrast, social understandings of disability, such as the social model of disability and postconventional or critical disability studies (Shildrick 2012), do not equate impairment with disablement. Disability – that is, the disadvantage experienced by disabled people – is seen to arise from unjust social arrangements and ableist ideologies. Impairment is bracketed and neither the social model or postconventional disability studies say much about the normative issues surrounding impairment; their role in promoting or hindering well-being and people's pursuit of a good life (Shakespeare 2014; Vehmas 2004; Vehmas and Watson 2014). The social model has concentrated on the societal and material causes of disability whereas postconventional disability studies has mainly produced genealogies and cultural analyses exploring the origins of ableist, discriminatory and oppressive ideas and values. Whilst both of these accounts have a normative motivation in that they seek to promote a society in which the disadvantage experienced by disabled people is removed, the normative dimension in the approaches goes virtually unacknowledged.

It is our aim in this article to address this lacuna and discuss the conditions for the normative scrutiny for disability studies. The first step of establishing such scrutiny is to define what normativity is and to rectify the confusion that currently exists in the way disability studies (especially postconventional accounts) conceptualises normativity and the concept of 'normative'. In postconventional disability studies the concepts 'normative' and 'normativity' have become conflated with 'normal' or 'normate', a conflation that is both unnecessary and false. 'Normative' and 'normality' do not necessarily relate to each other, and making the false comparison between the two may serve to restrict or even prevent fruitful evaluative discussion on important ethical and political issues that relate to disabled people's lives. The article then moves on to present some methodological considerations necessary for examining normative issues.

We suggest that the best way to deal with such issues is to resort to the methods provided by philosophical ethics and we propose a conventional analytic philosophical method to address ethical issues related to disability. After this, the article will discuss briefly two examples in postconventional accounts where the conceptual confusions and inadvertent application of theoretical ideas have resulted in confused normative judgments. In particular we will focus on defining agency and responsibility, and the recent posthumanist turn in disability studies and its pursuit to nullify the separation between humans and animals. We will argue that whilst these theoretical openings are interesting, they contain questionable elements which, without further explication, may unintentionally risk the moral agency and rights of some disabled people.

Our aim is not to provide an argument regarding responsibility or moral status (that is beyond the scope of this article) but, rather, to unpack the problematic relationship postconventional approaches have with normativity and its implications. In the final section we offer some reflections on a crucial, related matter; namely, the use of empirical claims in making normative arguments. We will argue that selective and simplified use of empirical knowledge

easily misconstrues the normative issues under discussion. Making normative analyses and judgments related to disability cannot be disconnected from the lived experiences of disability. This implies that sound normative analysis requires a sound use of empirical data.

What is 'normative' and what it is not

The usage of the term 'normative' or 'normativity' varies in disability studies. Postmodern formulations have confused the issue by attaching unconventional meanings to the concept, and it has joined Enlightenment as almost a term of abuse. Smith (2004), for example, seeks to oppose definitions of disability based on what he calls 'normative, positivist, monosemic, professionally-mapped thing founded in disease and filth metaphors'. Garland-Thomson describes how, for her, disabled people are employed as 'a synecdoche for all forms that culture deems non-normative' (2013, 335). Normativity and the normate are seen as products of contemporary post-enlightenment science (Dolmage 2014) and as a major source of the oppression faced by disabled people. Because of this, Goodley (2014, 158) calls for what he terms a non/normative dis/ability studies, and states that 'disability demands non-normative and anti-establishment ways of living life' (Goodley, Lawthom, and Runswick-Cole 2014, 348). The concept 'normative' thus is often used confusingly and has become a synonym for 'normal', 'normate' or 'standard' in disability studies.

This usage of the concept 'normative' is misguided in these accounts and it unwarrantedly confuses and complicates the issue. Normativity is an inescapable part of social life – it pervades our lives:

> We do not merely have beliefs: we claim that we and others ought to hold certain beliefs. We do not merely have desires: we claim that we and others not only ought to act on some of them, but not others. (O'Neill 1996, xi)

Meaningful dialogue requires mutual agreement of the meaning of words, concepts, laws and so forth; it requires a normative agreement and commitment that a particular word, for example, has one meaning rather than another or that a particular law either condones or judges some actions:

> … the successful performance of any speech act presupposes norms of truth, comprehensibility, truthfulness, and appropriateness. Such norms make communication possible, but only by devaluing and ruling out some possible and actual utterances. They are *what enable* us to speak, at the same time and insofar as they *constrain* us. (Fraser 1981, 285; original emphases)

On the basis of such normative agreements, we are able to judge statements as correct or incorrect, just or unjust, right or wrong, valid or invalid, and so on. In other words, the normative is a special realm of fact that validates, justifies, makes possible and regulates rules, meanings and reasoning. Where there are humans and social life, there is necessarily normativity as well (Fraser 1981, 285; Sayer 2011, 23; Turner 2010, 1–2).

This is the way the concept 'normative' is conventionally used in philosophy; normative judgments, properties, propositions, facts and the like are those that fall on the ought side of the is–ought distinction and on the value side of the fact–value distinction (Enoch 2011, 2–3). Paradigmatic examples of normative propositions would be 'we ought to support gender equality', 'women have a moral right to practice procreative autonomy' and so on. There are, of course, controversies about how best to make theoretical sense of the 'fact–value distinction' or 'is–ought distinction', whether such distinctions can justifiably be made in the first place, and whether normativity is a feature of the world or feature of concepts and

words (Finlay 2010). However, we are not and we do not need to be concerned about these controversies in this article.

We use the term 'normative' here rather broadly to include all evaluative judgments of an ethical sort. We concentrate on a particular subset of normative propositions; namely, ethical judgments concerning disability (political statements are often also normative, but due to the lack of space we will concentrate on ethics). There are various alternative ethical frameworks (e.g. religions) to address issues of moral concern but, in our view, the best way to approach ethical issues is to resort to the analytic tools provided by philosophical ethics.

The field of philosophical ethics is usually divided into three subject areas: meta-ethics, normative ethics and applied ethics. Meta-ethics examines the origin and meaning of ethical concepts covering metaphysical issues such as whether morality exists independently of human beings, and psychological issues about the mental basis of moral conduct and the motivation to be moral (for example, Sayre-McCord 2014). Normative ethics is the general study of goodness and right action that aims to formulate moral standards which regulate right and wrong conduct. The main questions of normative ethics are: 'What kinds of beings should we be like?' and 'How are we to live?' Normative ethics thus aims to describe the best features of human character and manner in a way that could be the basis for normative rules and even law-making and jurisdiction that guide human conduct and behaviour (Frankena 1973; Timmons 2002, 3–4). Applied ethics, in its part, analyses specific, often controversial, practical issues such as abortion, euthanasia or disability rights – the kinds of issues that are largely seen to have a special moral significance. One characteristic of ethical norms regarding cases of special social importance is their compelling and even overriding nature compared with other norms; morality concerns universally or generally obligatory practices (Hare 1963; Scanlon 2000). Moral norms thus may override legal norms should they happen to conflict with moral sentiments. For example, homosexuality was until relatively recently illegal in many western countries but was finally legalised due to the changed general moral convictions. Jurisdiction that conflicts with morality lacks credibility and authority, and people in general are reluctant to knowingly support initiatives or policies that they see to be morally wrong.

Normativity as an inescapable dimension of disability

Disability studies has always included a strong normative dimension, founded as it is on a belief that life for disabled people could be better coupled with a desire to identify and challenge what are seen as discriminatory practices and beliefs. All theoretical accounts in the field contain either implicit or explicit normative judgments about the ethical or political issues that affect disabled people's lives. Without considering properly the normative dimension related to disability, disability scholars cannot fulfil the practical and theoretical aim of disability studies, namely: to understand disability better, to develop and design appropriate policy responses, and, in general, to make things better for disabled people (Vehmas and Watson 2014).

Postconventional disability studies is not keen to make normative arguments and judgments. This does not, however, mean that it does not adopt a normative stance; it is after all opposed to the discrimination experienced by disabled people, seeks to promote disabled people's well-being and improve their life opportunities. This is an example of 'crypto-normativity' where the normative judgments and their reasons are hidden, and

the arguments are presented as if they were non-normative (Fraser 1981). This provides an incomplete and inadequate analysis. To paraphrase Fraser (1981), by rejecting normative scrutiny and the production of clearly articulated normative judgments, postconventional disability studies presumes that we all already agree about what it is that we are opposed to – discrimination against disabled people, hate crime, denial of opportunity, hierarchical treatment of disabled people and so forth – and also that we all agree on why they are bad (they deny opportunity, they remove rights, etc.). Normativity thus becomes a series of 'oughts' with little or no evaluation as to why they are classified as such (Sayer 2011). By taking this approach, postconventional disability scholars can no longer account for their own explicit normative judgment, as Fraser (1996, 216) argues in her critique of Butler.

Scholars working in this intellectual tradition typically insinuate how things, as they currently stand, are wrong whilst providing very little practical ethical guidance as to how things ought to be (Sayer 2011, 44 and 229; Vehmas and Watson 2014). Whilst postconventional disability studies clearly is an ethical and political project, exemplified in many recent works (for example, Goodley, Lawthom, and Runswick-Cole 2014; Shildrick 2012), these accounts lack an engagement with the normative issues of how things ought to be. Current postconventional normative analysis does not properly articulate the moral and political wrongs disabled people face, and neither does it give us practical guidelines regarding ways to right those wrongs (cf. Lundblad 2011, 146).

Goodley and Runswick-Cole (2015), for example, have recently produced an analysis of Big Society ideas in the United Kingdom and how the present neo-liberal agenda has damaged disability services and disabled people's social status in general in the United Kingdom. They conclude that we should question ableist norms guiding social arrangements, and that we should 'continue to question, destabilise assumptions that marginalise and exclude bodies and minds that are judged to fail to meet the expectations of ableist normativity' (Goodley and Runswick-Cole 2015, 10). Using the schema developed by Fraser (2013), we would argue that there are two clear claims in the arguments contained in this and other works in postconventional disability studies (see, for example, McRuer 2013; Wolbring 2008). Firstly, disabled people emerge via regulatory systems and forms of normality, and it is only in these highly ableist regulatory schemas that the contemporary disabled subject is constructed. Disabled people become constituted through the power/discourse formations where all subjectivities are always already culturally constructed (Fraser 2013). There is in essence no doer behind the deed.

Secondly, by drawing on the ideas of Foucault, Butler and other poststructuralists, postconventional disability studies implies that the normality-driven practices of subjectification drive subjection. That is, the disabled subject emerges through normative practices that prioritise normality, which serves to exclude and silence 'the disavowing quality of the normal' (Goodley 2014, 157). It is only through the silencing of others that authorised subjects can emerge, or, as Goodley (2014, 166) puts it, production of the Über-able results in the production of the ultra-disabled.

This is all fine and good but what does this mean in practice? What kind of guidance does it offer in relation, for example, to impairment-related violence, or improving access to goods and services for those who harm themselves or who are at risk? What does this guiding principle imply in terms of social policy or how to best arrange our societies in order to meet people's needs fairly? We are fully with Goodley and Runswick-Cole (2015, 10) when they argue that 'theory can help us to create opportunities for the urgent acts of

refusal, revision and resistance needed to bring people in from the cold', but their words leave us in bafflement as they fail to explicate or offer concrete, constructive examples of how to refuse, revise and resist, and on what grounds. If disability is reduced to discursive norms it becomes difficult to see how policies (always being already normative) can emerge. There is little analysis within postconventional disability studies that explores the economic processes, the social relational and organisational life and the material interests and other non-discursive forms of power that create disability.[1]

Here we may be accused of committing the 'fallacy of externalism', meaning that we criticise postconventional disability studies for failing to accomplish something which these theorists did not set out to pursue in the first place (Baert and Carreira da Silva 2010, 9). Is it not the case that postconventional disability studies mainly analyses and describes how certain ideas came about without even attempting to offer any normative guidance of how things ought to be? So is our criticism misguided? No, because postconventional accounts of disability highlight ableist ideas, ideologies, representations, policies and so forth that create and further disabled people's exclusion, marginalisation and oppression. Thus, these accounts point out moral and social evils which means that normativity is at the heart of their analyses. This being the case, it is only proper to ask whether their theories provide us with any tools to settle the moral and political wrongs they have raised.

Methodological considerations

If we want to say something constructive and valid about whether certain practices or policies are morally right or wrong, we need appropriate analytic tools. Producing sound normative arguments regarding issues to do with, say, health, autonomy, dying and justice is admittedly very difficult. Whilst normative issues cannot be settled once and for all with simple knock-down arguments, it is possible to generate constructive normative arguments – a task that we should not give up entirely to philosophers, who often have very little awareness of the complexities of disability. We are highly sceptical whether poststructuralism offers sufficient tools to unpack such issues properly.

We will now outline briefly an application of the traditional analytic philosophical method which, in our view, provides a useful starting point for addressing ethical issues related to disability (see Tooley 1983, 11–18).[2] The first thing to do is to 'clarify and evaluate statements and terms that are vague or ambiguous'. This point expresses an aspiration for conceptual clarity and logical soundness. For example, some prominent philosophers have argued that intellectual disabilities are relevant factors when considering abortion or infanticide and they have appealed to such vague and ambiguous concepts as 'recognizably human life' (Rachels 1986, 66), 'normal life' (Kuhse and Singer 1985, 143) and 'full human existence' (Kuhse and Singer 1985, 136). Clearly, these kinds of utterances express ideas of what kind of beings human beings ought to be and what kinds of lives they ought to live, but at the same time they are too vague to be used as a basis for normative judgments regarding, for example, end-of-life issues (Vehmas 1999). Second is 'detecting and rejecting unreasonable factual claims of a non-ethical sort'. The use of factual, empirical claims inevitably directs the normative arguments and their conclusions. Thus, in the case of disability it is important to acknowledge in a balanced way relevant environmental (including social and geographical) circumstances, historical and cultural factors as well as individual factors (such as impairment effects) that ultimately give meaning to the phenomenon of disability. Third is 'detecting

inconsistencies in ethical positions'. Avoiding logical confusions is a prerequisite to the production of sound and valid (philosophical) argumentation (for example, Häyry 1994, 147–158; Tooley 1983, 11–18). Applying this tool to the examination of postconventional accounts of disability is challenging because of their opacity, even obscurity. We are not always certain what the normative premises of their arguments are, let alone the substance and the logic of their arguments.

One traditional way to produce normative arguments on practical moral problems is to apply an ethical theory to the issue under discussion. One could adopt Kantian premises as a basis for argumentation where certain principles as regards respecting humanity as an end itself, and the intentions of the agents, would be the central points determining the morality of actions (for example, Baron 1997). Or one could resort to consequentialism, which emphasises the obligation to do whatever will produce the best consequences, or the obligation to avoid creating bad consequences (for example, Pettit 1997). A third central approach to ethics is virtue ethics, an agent-focused account with the focus on the virtuous individual and on those inner traits, dispositions and motives that qualify her as being virtuous (for example, Slote 1997). Ethical theories give useful tools and elements for argumentation in applied ethics, but using merely one ethical theory as the criterion for moral judgments can result in simplified, one-sided arguments and views that fail to consider the complexity and variety of moral realm related to disability (a classic example of this would be Singer's writing on disability; see later).

One fruitful way of unpacking, for example, the significance of impairment regarding well-being and a good life is to conceptualise the issue in terms of instrumental and intrinsic considerations and whether disadvantages related to disability are the result of impairment effects, disablism or a combination of both. Instrumental factors are those things that enhance human well-being, and provide the means to achieve things that are of intrinsic value. Intrinsic factors, on the other hand, are things that are intrinsically valuable; that is, valuable in a way that needs no further explanation or justification (such as happiness). These criteria are problematic in many ways but they are also inevitable: we do need to make judgments about what is good or bad for people in order to make decisions regarding, for example, medical treatment, public policy or education (Vehmas 2012; Vehmas and Watson 2014).

In the following sections we will discuss two issues recently raised within postconventional disability studies; namely, suggestions about new ways to conceptualise agency and responsibility, and human–animal relation. We have selected these examples because they contain normative elements which the authors of these accounts have ignored. This shortcoming is a logical outcome of the postconventional disability studies that plays with novel ideas without proper engagement with their normative implications. In what follows we employ these concepts to examine the soundness of these suggestions and their validity in making normative judgments. We will point out that these conceptualisations, whilst interesting, are also ambiguous and fail to take into account the following normative corollaries.

Responsibility and agency: ableist ideals of enlightenment?

The notions of competency and capability have long been used as a means of excluding disabled people from a range of different settings including education, work and leisure activities, as well as denying them access to opportunities such as relationships, parenting

and friendships (Goodley 2014, 156). The response of much of postconventional disability studies has been to analyse the cultural origins of this exclusion, but whilst this genealogy may explain why this has happened and explore the ideological origins, it fails to unpack carefully why they are wrong. Goodley acknowledges that there is a need for a normative stance but argues that calls for normative judgments run the risk of 'leaving dis/ability only in the realm of normative [*sic*] … [and] ignoring the productive potential of dis/ability to crip a whole host of norms that are, in actuality, limiting and stifling' (2014, 158). These, he suggests, include characteristics such as rationality, capability, responsibility and competency; which he identifies as ableist ideals, part of the modernist humanistic subject (2014, 156). This raises a number of questions such as whether rationality and responsibility are normatively on a par with irrationality and irresponsibility. Surely not. Homophobia, for example, is an irrational and morally harmful mindset. It is based on irrational fears about contagion, divine retribution and so forth, and results in moral harms to sexual minorities such as discrimination, hate speech and hate crime.

How about (ir)responsibility? Imagine Bill is by himself taking care of his infant child at his family's remote country house and all of a sudden feels an irresistible urge to drink three bottles of wine. Getting intoxicated is, perhaps, a perfectly legitimate interest as such, but in those circumstances it would compromise Bill's capability to properly supervise his child and thus put the child at risk and in danger. In other words, rationality and responsibility in general are good, irrationality and irresponsibility are bad.

These ideals as such are not the problem; the problem is their application and the way they are used to discriminate against disabled people. For example, parents with intellectual disabilities are too easily labelled as incompetent and irresponsible because of their cognitive capacities, as Goodley (2014, 156) points out. But judging whether someone is fit to raise children does not necessarily need to involve disablism of any kind. Parents who neglect or maltreat their children for whatever reason may do so despite excellent support. In these kinds of undoubtedly very difficult cases, society has not just a right to intervene and take their children into custody, but actually a duty. Things like child protection as such are not about, for example, disability or drug abuse, they are about promoting children's well-being (we are, of course, fully aware that in practice child protection does too often include disablism and other forms of prejudice). Thus, we should be careful not to throw out the baby with the water and we should not abandon a perfectly legitimate ideal due its occasional misplaced application. Removing disablism or ableism from the equation will not resolve all ethical issues related to disability.

Goodley, Lawthom, and Runswick-Cole (2014) argue that we should rethink the humanist notions of agency, responsibility and subjectivity, and they have drawn on Braidotti's posthuman subjectivity, which argues for 'partial form of accountability, based on a strong sense of collectivity, relationality … an affirmative bond that locates the subject in the flow of relations with multiple others' (Braidotti 2013, 49–50). These kinds of formulations with references to partial accountability raise questions regarding moral and legal responsibility: should we abandon the idea of an individual as an agent who is responsible for his or her actions? It is not clear what Braidotti or Goodley et al. specifically imply with these kinds of formulations. However, their calls for posthuman subjectivity can, if consistently applied, be seen to include notions of collective agency and responsibility. Assuming this is the case, we will now briefly examine whether this would be a reasonable alternative foundation in understanding subjectivity, agency and responsibility.

Western philosophy and our legal system are based on the premise that we can reasonably ascribe praise or blame to autonomous people for their actions and omissions. Responsibility ascriptions require moral agency, meaning, the capacity to evaluate reasons for acting. Consider, for example, that your careless cat accidently breaks your valuable ceramic sculpture. Whilst the cat was causally responsible for the destruction and may well make you feel regret and even anger, to feel moral indignation would be clearly unwarranted. The cat is not a moral agent, a being that is capable of acting with reference to right and wrong. On the other hand, if the sculpture was deliberately broken by a neighbour who was envious and broke it in order to hurt you, resentment and moral indignation would be appropriate. In other words, judgments about individuals' moral responsibility relate to views about, first, their relevant capacities to evaluate reasons for acting, second, their possibilities to act freely and, third, possible excusing factors (Eshleman 2009; Fischer and Ravizza 1998, 1–2). All of these three factors are relevant to disability studies and are unavoidable when considering praiseworthiness or blameworthiness. For example, what kinds of capacities are relevant for one's responsibility and how do we measure them? Or what kinds of conditions may limit one's capacity to act freely and in ways that are out of their control; are the acts of someone with, say, antisocial personality disorder merely involuntary reactions that stem from his or her brain functions – a bit like sneezes or seizures and thus forgivable (Arpaly 2005, 290–291)?

Issues to do with moral responsibility are complex but they are even more complex when responsibility is extended to groups: can we meaningfully talk about collective responsibility? Collective responsibility means ascribing causal responsibility and blameworthiness to groups and their collective actions. Ideas about collective agency and responsibility raise various difficult issues such as whether collective responsibility is an aggregative phenomenon, something that transcends the contributions of individual group members. Imagine several people overturning a police car in a protest; is it more accurate to describe this act in terms of shared responsibility where each individual taking part in the action is individually responsible for the state of affairs, or should we assign responsibility to a single entity, the collective, consisting of the number of people who constitute it? A related matter is that of agency. Moral responsibility can reasonably be seen as something that requires agency; that is, capability to act intentionally. If we assume the simple view that an 'agent is, by definition, something that acts; and if, at a time, something acts, that thing is an agent then' (Mäkelä 2013, 8), can collectives be seen as agents? It would seem peculiar to claim that groups think as groups or that they formulate intentions (i.e. mental states) that guide their actions. It seems more credible to think that it is the persons who constitute the group, and it is their individual intentions and actions together that form collective actions. But, on the other hand, there are cases that cannot be true of individuals but only of collectives; for example, a football match is won by a team and a president is elected by a nation. It is only collectives that can perform such acts and these acts are based on collective beliefs, intentions and agreements. At the same time, however, it would seem odd to infer that this proves the existence of a collective mind which guides the group's actions. Instead, the group members' individual minds produce beliefs, desires and so forth, and together they form a kind of derivative collective mind with concomitant collective actions. Even if it was correct to ascribe causal responsibility and blameworthiness, and thus moral responsibility, to a collective entity, various worries would remain; would collective responsibility have any moral, judgmental force and, if it did, how could we avoid injustices where agents were held responsible for someone else's actions?

Talk about collective or shared agency and responsibility is by no means arbitrary in the case of, for example, ethnic communities and social movements, but it does include problems that should not be dismissed lightly in disability studies. If we are to apply these ideas to disability, we have to specify the ontology of such a conception (what is a collective agent as opposed to an individual agent) as well as the ethical issues related to it (how can we ascribe moral responsibility to groups). The idea of 'expanded identity' (Goodley and Rapley 2002) seems useful and plausible in the case of, for example, people with intellectual disabilities, but it should be implemented with caution because if we absolve someone from individual moral responsibility and individual moral agency, we might be seen to question and nullify that individual's entitlement to moral rights.

Posthumanism and disability: why being human matters

Western philosophy and thought in general have traditionally emphasised certain characteristics that separate humans from non-human animals and what makes us distinctively human. In philosophy, the concept of a 'moral being' denotes a being that merits rights and is capable of acting morally. It has been argued that only rational beings can be moral, which allegedly implies that non-human animals as well as some people with cognitive disabilities are 'amoral' beings. Their behaviour, even if harmful or otherwise undesirable, is not seen as immoral but merely as regrettable, an unavoidable consequence of their not knowing better. In everyday morality and in moral philosophy (at least in the works of such classics as Plato, Aristotle, Kant and Mill), rationality and practical reason are attributed to adult human beings whose intelligence is 'normal' (Sapontzis 1980; Vehmas 2004). Thus, it can reasonably be argued that many disabled people are, according to some ethical theories, marginal human beings (Silvers 1998, 3). This moral philosophical tradition is found in contemporary philosophical ethics by such notable philosophers as Peter Singer (1993) and Jeff McMahan (2002).

Understandably, many disability activists and researchers have had a critical and hostile attitude towards philosophy and philosophers (Oliver 2007; Schöne-Seifert and Rippe 1991). Speculations about the moral significance of individual characteristics have been abandoned as disablist and even eugenic (Koch 2011), and the mere humanity of all disabled people has been seen as the axiomatic foundation of their equal, unquestionable moral worth. Considering this history, it is somewhat surprising that posthumanist accounts have gained increasing popularity in disability studies (Goodley, Lawthom, and Runswick-Cole 2014; Nayar 2013).

Posthumanism argues, among other things, that human beings should not have *a priori* ethical primacy over non-human animals, and that any moral hierarchy as well as the division between humans and animals is false and based on 'ethical parochialism' (Wolfe 2010, 61). Goodley, Lawthom, and Runswick-Cole have used Braidotti's ideas of contesting anthropocentrism and the superior moral value of humans compared with other species, and valuing death as much as life, to claim that:

> disability is *the* quintessential posthuman condition: because it calls for new ontologies, ways of relating, living and dying. (Goodley, Lawthom, and Runswick-Cole 2014, 348; original emphasis)

This contention is at best vague and equivocal and creates more questions than it answers.

If the humanist legacy and its conviction of the superior moral value of human beings over non-human animals are abandoned, postconventional disability studies ends up in

the murky theoretical waters of moral status. This is clearly something they do not want to do, probably because their poststructuralist theoretical tools give them very little capacity to deal with such issues (see Sayer 2011). Imagine, for example, that you can save from a burning building either a human that is a stranger to you or your beloved dog. Should a paramount moral consideration in such a case be the biological species, or the presumably more advanced cognitive capacities of the human, or the special relation you have to your dog?

The most well-known and carefully articulated arguments with the aim to abolish anthropocentrism and speciesism in ethical thinking are those provided by philosophers promoting animal rights. Both Singer (1993) and McMahan (2002) have offered arguments to support the notion that belonging to the species *Homo sapiens* is morally irrelevant. What matters morally, they argue, are intrinsic psychological characteristics and other morally significant factors such as sentience. Based on such reassessments of moral status, being human no longer automatically secures one's moral worth but, rather, their psychological and intellectual capacities.

It is easy to theorise that humans and animals form an assemblage of subjects with equal worth, but it is more difficult to put this ideal into practice. In a world with limited resources, should we prioritise, for example, the health care of animals rather than humans? If not, clearly humans are morally more valuable than animals. If one finds this widely held moral intuition unacceptable, one should come up with an argument that plausibly proves that moral worth should be distributed equally to different species, and accept the logical conclusions of such an argument. Toying with the ideas provided by posthumanism may be useful to disability studies, but in our view such ideas should not be adopted without caution. For example, what would be the implications of 'the displacement of anthropocentrism and the recognition of trans-species solidarity' and viewing 'subjectivity as an assemblage that includes non-human agents' (Braidotti 2013, 67 and 82)? Goodley, Lawthom, and Runswick-Cole seem to suggest that it is not problematic to make a comparison between non-human animals and disabled people, because all humans are ultimately animals too:

> The problem is not that some categories of human are treated like animals; the problem resides in the unconscious desire of the human condition to treat animals in inhumane ways; and treat some humans as if they were animals. We think that reinvigorating discussion around human/animal relations around disability might provide the necessary conditions and impetus for revaluing animals and humans as sharing a posthuman space of becoming. (2014, 355)

This call for revaluing animals and humans as constituting a common posthuman condition somewhat resembles, again, the points raised by philosophers supporting animal rights and questioning the equal worth of humans with, for example, profound intellectual disabilities. We assume that posthumanist disability scholars do not wish to join Singer and McMahan in their ableist and disablist endeavour, and if so perhaps they should be a bit more hesitant in their embrace of Braidotti's ideas. If we are to do away with categories and moral hierarchies between different biological species, and hold on to the idea of humans and non-humans constituting an equal moral class, one ends up in highly counter-intuitive conclusions (Curtis and Vehmas 2016). In this ethical scenario, it would be wrong to assume that human beings in themselves are morally more valuable than pet dogs, and pet dogs are more valuable to their owners (and possibly even in themselves) than vertebrates in general, and that vertebrates are more valuable than non-vertebrates. These kinds of hierarchies result from virtually all traditional ethical theories, albeit for different reasons (Curtis and Vehmas 2014). If these hierarchies are to be replaced with the notion of 'posthuman space of becoming', things such

as sentience, capacity to reason and close relationships would perhaps no longer be morally paramount. But if that is the case, posthumanists need to argue in more detail exactly what should be the basis of the moral worth of different kinds of humans and animals.

Further, consider McMahan's argument that those diagnosed with severe cognitive disability are psychologically comparable with non-human animals such as dogs, and are only able to achieve a level of well-being equal to that of 'a contented dog' (2002, 153). Most people would find this comparison at best questionable, if not offensive. This is because our collectively held ethical convictions are constructed narratively due to the various historical happenings and processes that have contributed to the general conceptions of right and wrong, good and bad. Particular historical episodes, such as the holocaust or the systematic mistreatment of disabled people in twentieth-century institutions, have had a tremendous effect on our moral thinking and at least in principle we consciously aim to prevent such atrocities ever happening again. One upshot of this is the general belief that we should pay special attention to protecting the equal worth of various minorities, including individuals with profound cognitive disabilities. In other words, regarding such individuals as morally more valuable than non-human animals is a reasonable upshot of the collective moral narrative of the western world. Naturally, this does not prove that maintaining such narrative ethical norms is philosophically justified, but pragmatically, and from the viewpoint of humans with profound cognitive disabilities and their families, it makes perfect sense. History has also taught us that it makes perfect sense to shun philosophical analyses that draw a parallel between individuals with such disabilities and animals – regardless of the empirical validity of the equation. Thus, perhaps there are, after all, good normative reasons to separate non-human animals from humans.

The humanist tendency to favour human beings compared with other beings can be seen as a prejudice similar to racism and sexism, or it can be seen as an inevitable part of morality being fundamentally a human endeavour; it is, after all, 'closely tied to the human perspective and the human motivational capacity because its point is the regulation of human conduct' (Nagel 1986, 186). Besides, by definition, being humane involves not only showing compassion and tenderness towards humans, but to non-human animals as well (we do not, however, wish to contradict the claim that humans do treat animals systematically in inhumane ways) (Williams 2006, 147).

All in all, posthumanism may provide useful insights and we agree that animality should be seen as a crucial part of humanity. This would not be a new idea; it would simply be a reformulation of Aristotle's notion of humans as political animals. In this view, rationality, sociability and animality are aspects of humanity that are thoroughly unified. Our animality includes corporeality and the inevitable dependency on others and their care for us in the different phases of our lives (MacIntyre 1999, 81–98; Nussbaum 2006, 159–160). Also, perhaps we should reconsider the moral status of animals, how we treat them and live with them; for example, perhaps we should stop eating meat and grant at least some animals moral rights. But this should be kept separate from the moral worth of humans, and especially those humans who are in the most vulnerable position. Included here are disabled people, too many of whom are still forced to live in animal-like conditions. The superior moral worth of all human beings in comparison with animals is a widely accepted conviction that should not be lightly abandoned. Thus, we are not convinced that the posthumanist agenda to eradicate epistemic and moral division between humans and animals serves disability theory or the good of disabled people. Putting humans and animals in the same box opens up the door for speculations about moral worth where disabled people are likely to be on the losing side.[3]

Whilst there are good ethical reasons to reconsider the moral worth of animals, one should, however, be cautious to suggest the reconsideration of the dichotomy between human beings and animals, let alone the moral supremacy of human beings. The requirement that ethical decision-making should be made from an impartial viewpoint to consider all beings equally is impossible if it is not 'involved in the peculiarities of the human enterprise' (Williams 2006, 147). Bernard Williams (2006, 148) argues that this is the unavoidable consequence of the evolutionary plateau we live on at the moment, and the range of abilities it provides us with. In the disability context this makes perfect sense; we live in a disablist world that continues to question the moral worth of disabled people. When we pursue theoretical explanations that would reduce or hopefully remove this form of oppression, we have to take into account the reality where it is implemented. The reality, in our view, is still too ignorant and dismissing about disability rights in order to suggest comparisons between disabled human beings and non-human animals.

Conclusion: empirical claims and normativity

Philosophical ethics, particularly bioethics, has been criticised for construing disability in a reductionist manner and downplaying the social factors affecting people's disablement. Disability is too often portrayed in a homogeneous, uncontextualised, uninformed and prejudiced way where the complexities of the lived experiences are ignored (Vehmas 2004, 2012). In short, philosophers have not let the empirical reality get in the way of argument. We are concerned, however, that the emphasis of the cultural at the expense of the material found in postconventional disability studies has a similar kind of tendency to emphasise one-sidedly only certain characteristics of disability whilst downplaying others, and this runs the risk of misconstruing not just the lives of disabled people but also the crucial ethical issues.

It is important to recognise the fine line between descriptive and prescriptive empirical claims. For example, to describe disability in terms of opportunity (Goodley 2014) or limited opportunity (Harris 2001) actually prescribes the empirical description either in a positive or negative fashion; Goodley (2014, 158) makes an explicit normative judgment by describing the lived experience of impairment in terms of 'productive potential' or 'cripping potentialities', whereas Harris (2001, 384) interprets similar empirical reality as a deprivation of possibilities. Bearing this in mind, it is crucial to pay attention to the way empirical data are used. We would suggest that there are at least three factors that should be born in mind when using empirical claims to support normative judgments. First, it is crucial to recognise the societal and cultural (usually disablist) reality where disability is defined and experienced. The surrounding social, cultural and ideological environment necessarily affects the subjective experiences of disablement. Secondly, as Thomas (1999) reminds us, impairments matter. The surrounding community, society and natural environment can make a dramatic difference in terms of impairment effect but, at the end of the day, one's life in a society, in a culture, is an embodied existence which affects the way one deals with the environment and with others. This notwithstanding, we need to consider carefully the impact of different impairments in different circumstances and how subjective experiences reflect historical and cultural consciousness. Nevertheless, the difficulties that disabled people confront are not only matters of access to resources and public spaces – they are cultural and aesthetic as well, but importantly they are always embodied (Blume 2012, 354; Shakespeare and Watson 2001). For example, chronic pain can greatly affect disabled children's subjective well-being (Colver et al. 2015). Fatigue, the gradual loss of muscular functioning, loss of memory, sight

or hearing are central to the experience of impairment effects to many disabled people, and to ignore these experiences would be one way to produce skewed accounts of disability that fail to do justice to disabled people's experiences. Thus, 'the corporeal quality of disabled experience [should be] reclaimed', as Stuart Blume (2012, 354) has put the matter. Giving more room to subjective voices will challenge theory, making it uncomfortably unpredictable as they may not submit to single theoretical accounts that emphasise the role of either material, cultural or structural in their constructions of disablement.

Third, empirical representations have normative implications. There is no absolute neutral ground, and because disability studies engages directly with those who suffer moral and political wrongs, research must be guided by a moral and political commitment to disability rights (see Oliver 1992). This means that whilst we can and should emphasise the positive potentialities of disability, we must not ignore or downplay the undesirable effects impairments sometimes have. Acknowledging the negative elements related to some impairments does not conflict with celebrating difference. It is only sensible and sensitive, and is required in order to ensure proper care and treatment. Our normative conclusion, thus, is that disability studies needs to maintain a normative commitment to benefit the well-being of disabled people without rejecting truthful empirical descriptions, and risking the integrity of academic research.

Notes

1. Similar critiques have been applied to poststructuralist accounts of gender by Connell (2012).
2. This methodological outline is admittedly cursory. Owing to space limitations we cannot present various requirements related to philosophical argumentation regarding, for example, the role of thought experiments, common sense, scientific evidence, generalisations and so on (for example, Daly 2010; Walton 2006).
3. We admit that some forms of posthumanism may provide a positive potential for the re-evaluation of the moral worth of human beings and non-human animals. As Wolfe (2010) argues, posthumanism need not reject humanism and the many valuable ethical notions it has introduced, but it needs to be critical to the ideological, ethical and political conventions of liberal humanism that exclude non-human animals and many disabled people from moral worth. To paraphrase Wolfe (2010, 137), we ought not to succumb to the kind of pragmatism that achieves certain gains in the short run, but at the price of a radical foreshortening of a more profound pluralistic ethical project that acknowledges the value of (human) beings other than white, able-bodied, heterosexual and so on.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by Kulttuurin ja Yhteiskunnan Tutkimuksen Toimikunta [grant number SA 275988].

References

Arpaly, N. 2005. "How is It Not 'Just like Diabetes': Mental Disorders and the Moral Psychologist." *Philosophical Issues* 15 (1): 282–298.

Baert, P., and F. Carreira da Silva. 2010. *Social Theory in the Twentieth Century and beyond*. 2nd ed. Cambridge: Polity.
Baron, M. 1997. "Kantian Ethics." In *Three Methods of Ethics: A Debate,* edited by M. W. Baron, P. Pettit and M. Slote, 3–91. Oxford: Blackwell.
Blume, S. 2012. "What Can the Study of Science and Technology Tell Us about Disability?" In *Routledge Handbook of Disability Studies*, edited by N. Watson, A. Roulstone, and C. Thomas, 348–359. London: Routledge.
Braidotti, R. 2013. *The Posthuman*. London: Polity.
Colver, A., M. Rapp, N. Eisemann, V. Ehlinger, U. Thyen, H. O. Dickinson, J. Parkes, et al. 2015. "Self-Reported Quality of Life of Adolescents with Cerebral Palsy: A Cross-Sectional and Longitudinal Analysis." *The Lancet* 385 (9969): 705–716.
Connell, R. 2012. "Gender, Health and Theory: Conceptualizing the Issue, in Local and World Perspective." *Social Science & Medicine* 74 (11): 1675–1683.
Curtis, B. L., and S. Vehmas. 2014. "Moral worth and Severe Intellectual Disability – A Hybrid View." In *Disability and the Good Human Life*, edited by J. Bickenbach, F. Felder, and B. Schmitz, 19–49. New York: Cambridge University Press.
Curtis, B. L., and S. Vehmas. 2016. "A Moorean Argument for the Full Moral Status of Those with Profound Intellectual Disability." *Journal of Medical Ethics* 42: 41–45. doi:10.1136/medethics-2015-102938.
Daly, C. 2010. *An Introduction to Philosophical Methods*. Peterborough: Broadview Press.
Dolmage, J. 2014. *Disability Rhetoric*. Syracuse: Syracuse University Press.
Eshleman, A. 2009. "Moral Responsibility." In *Stanford Encyclopedia of Philosophy*, edited by E. N. Zalta, Stanford: The Metaphysics Research Lab. http://plato.stanford.edu/entries/moral-responsibility.
Enoch, D. 2011. *Taking Morality Seriously: A Defence of Robust Realism*. Oxford: Oxford University Press.
Finlay, S. 2010. "Recent Work on Normativity." *Analysis* 70 (2): 331–346.
Fischer, J. M., and M. Ravizza. 1998. *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge: Cambridge University Press.
Frankena, W. K. 1973. *Ethics*. 2nd ed. Englewood Cliffs, N.J.: Prentice-Hall.
Fraser, N. 1981. "Foucault on Modern Power: Empirical Insights and Normative Confusions." *Praxis International* 3: 272–287.
Fraser, N. 1996. "False Anti-Thesis: A Response to Seyla Benhabib and Judith Butler." In *Justice Interruptus: Critical Reflections on 'Postsocialist' Condition*, edited by N. Fraser, 207–224. London: Routledge.
Fraser, N. 2013. *Fortunes of Feminism: From State-Managed Capitalism to Neoliberal Crisis*. New York: Verso.
Garland-Thomson, R. 2013. "Integrating Disability, Transforming Feminist Theory." In *The Disability Studies Reader*. 4th ed., edited by L. J. Davis, 333–353. New York: Routledge.
Goodley, D. 2014. *Dis/Ability Studies: Theorising Disablism and Ableism*. London: Routledge.
Goodley, D., R. Lawthom, and K. Runswick-Cole. 2014. "Posthuman Disability Studies." *Subjectivity* 7 (4): 342–361.
Goodley, D., and M. Rapley. 2002. "Changing the Subject: Postmodernity and People with 'Learning Difficulties'." In *Disability/Postmodernity: Embodying Disability Theory*, edited by M. Corker and T. Shakespeare, 127–142. London: Continuum.
Goodley, D., and K. Runswick-Cole. 2015. "'Big Society? Disabled People with the Label of Learning Disabilities and the Queer(Y)Ing of Civil Society." *Scandinavian Journal of Disability Research* 17 (1): 1–13.
Hare, R. M. 1963. *Freedom and Reason*. New York, NY: Oxford University Press.
Harris, J. 2001. "One Principle and Three Fallacies of Disability Studies." *Journal of Medical Ethics* 27 (6): 383–387.
Häyry, M. 1994. *Liberal Utilitarianism and Applied Ethics*. London: Routledge.
Koch, T. 2011. "Eugenics and the Genetic Challenge, Again: All Dressed up and Just Everywhere to Go." *Cambridge Quarterly of Healthcare Ethics* 20 (2): 191–203.
Kuhse, H., and P. Singer. 1985. *Should the Baby Live? The Problem of Handicapped Infants*. Oxford: Oxford University Press.
Lundblad, M. 2011. "Consider the Dragonfly: Cary Wolfe's Posthumanism." *Substance* 40 (3): 146–153.
MacIntyre, A. 1999. *Dependent Rational Animals: Why Human Beings Need the Virtues*. London: Duckworth.

Mäkelä, P. 2013. *Collective Responsibility: Against Collectivism*. Philosophical Studies from the University of Helsinki 40. Helsinki: University of Helsinki.
McMahan, J. 2002. *The Ethics of Killing*. Oxford: Oxford University Press.
McRuer, R. 2013. "Compulsory Able-Bodiedness and Queer/Disabled Existence." In *The Disability Studies Reader*. 4th ed., edited by L. J. Davis, 369–378. New York: Routledge.
Nagel, T. 1986. *The View from Nowhere*. New York, NY: Oxford University Press.
Nayar, P. K. 2013. *Posthumanism*. Cambridge: Polity.
Nussbaum, M. 2006. *Frontiers of Justice: Disability, Nationality, Species Membership*. Cambridge, MA: Harvard University Press.
O'Neill, O. 1996. "Introduction." In *The Sources of Normativity*, edited by C. Korsgaard, xi–xv. Oxford: Oxford University Press.
Oliver, M. 1992. "Changing the Social Relations of Research Production?" *Disability, Handicap & Society* 7 (2): 101–114.
Oliver, M. 2007. "Disability Rights and Wrongs?" *Disability & Society* 22 (2): 230–234.
Pettit, P. 1997. "The Consequentialist Perspective." In *Three Methods of Ethics: A Debate*, edited by M.W. Baron, P. Pettit, and M. Slote, 92–174. Oxford: Blackwell.
Rachels, J. 1986. *The End of Life: Euthanasia and Morality*. Oxford: Oxford University Press.
Sapontzis, S. F. 1980. "Are Animals Moral Beings?" *American Philosophical Quarterly* 17 (1): 45–52.
Sayer, A. 2011. *Why Things Matter to People: Social Science, Values and Ethical Life*. Cambridge: Cambridge University Press.
Sayre-McCord, Geoff. 2014. "Metaethics." *The Stanford Encyclopedia of Philosophy* (Summer 2014 Edition), edited by N. Edward Zalta. http://plato.stanford.edu/archives/sum2014/entries/metaethics/
Scanlon, M. 2000. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press.
Schöne-Seifert, B., and K.-P. Rippe. 1991. "Silencing the Singer: Antibioethics in Germany." *Hastings Center Report* 21 (6): 20–27.
Shakespeare, T. 2014. *Disability Rights and Wrongs Revisited*. 2nd ed. London: Routledge.
Shakespeare, T., and N. Watson. 2001. "The Social Model of Disability: An Outdated Ideology?" *Research in Social Science and Disability* 2: 9–28.
Shildrick, M. 2012. "Critical Disability Studies: Rethinking the Conventions for the Age of Postmodernity." In *Routledge Handbook of Disability Studies*, edited by N. Watson, A. Roulstone and C. Thomas, 30–41. London: Routledge.
Silvers, A. 1998. "Formal Justice." In *Disability, Difference, Discrimination: Perspectives on Justice in Bioethics and Public Policy*, by A. Silvers, D. Wasserman, and M. Mahowald, 1–145. Lanham, MD: Rowman and Littlefield.
Singer, P. 1993. *Practical Ethics*. 2nd ed. Cambridge: Cambridge University Press.
Slote, M. 1997. "Virtue Ethics." In *Three Methods of Ethics: A Debate*, edited by M.W. Baron, P. Pettit, and M. Slote, 175–238. Oxford: Blackwell.
Smith, Phil. 2004. "Whiteness, Normal Theory, and Disability Studies." *Disability Studies Quarterly* 24 (2). http://dsq-sds.org/article/view/491/668
Thomas, C. 1999. *Female Forms: Experiencing and Understanding Disability*. Buckingham: Open University Press.
Timmons, M. 2002. *Moral Theory: An Introduction*. Lanham: Rowman and Littlefield.
Tooley, M. 1983. *Abortion and Infanticide*. Oxford: Clarendon Press.
Turner, S. P. 2010. *Explaining the Normative*. Cambridge: Polity.
Vehmas, S. 1999. "Newborn Infants and the Moral Significance of Intellectual Disabilities." *Journal of the Association for Persons with Severe Handicaps* 24 (2): 111–121.
Vehmas, S. 2004. "Ethical Analysis of the Concept of Disability." *Mental Retardation* 42 (3): 209–222.
Vehmas, S. 2012. "What Can Philosophy Tell Us about Disability?" In *Routledge Handbook of Disability Studies*, edited by N. Watson, A. Roulstone and C. Thomas, 298–309. London: Routledge.
Vehmas, S., and N. Watson. 2014. "Moral Wrongs, Disadvantages, and Disability: A Critique of Critical Disability Studies." *Disability & Society* 29 (4): 638–650.
Walton, D. 2006. *Fundamentals of Critical Argumentation*. Cambridge: Cambridge University Press.
Williams, B. 2006. *Philosophy as a Humanistic Discipline*. Princeton: Princeton University Press.
Wolbring, G. 2008. "The Politics of Ableism." *Development* 51 (2): 252–258.
Wolfe, C. 2010. *What is Posthumanism?* Minneapolis, MN: University of Minnesota Press.

Australasian Philosophical Review

Volume 4 Issue 4 December 2020

Routledge
Taylor & Francis Group

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/rapr20

Rationality versus Normativity

John Broome

To cite this article: John Broome (2020) Rationality versus Normativity, Australasian Philosophical Review, 4:4, 293-311, DOI: 10.1080/24740500.2021.1964236

To link to this article: https://doi.org/10.1080/24740500.2021.1964236

Published online: 04 Nov 2021.

AUSTRALASIAN PHILOSOPHICAL REVIEW
2020, VOL. 4, NO. 4, 293–311
https://doi.org/10.1080/24740500.2021.1964236

Rationality versus Normativity

John Broome

University of Oxford and Australian National University

ABSTRACT

Philosophers often do not make as sharp a distinction as they should between rationality and normativity. Partly this is because the word 'reason' can be used to refer to either, and this leads to a confusion over meanings. This paper starts by clarifying the meanings of 'normativity' and 'rationality'. It argues that it is a conceptual truth that rationality supervenes on the mind. Then it considers substantive arguments that purport to show there is no real distinction between rationality and normativity. Many philosophers give a reductive account of rationality in terms of reasons. In particular, many claim that rationality consists in responding correctly to reasons. Since responding correctly to reasons is the concern of normativity, this in effect identifies rationality with normativity. This paper denies that rationality is identical to normativity, by means of what I call a 'quick objection'. The quick objection is that rationality supervenes on the mind whereas complying with normativity does not. I consider and reject some ways of responding to the quick objection, including an argument by Kiesewetter to the effect that normativity supervenes on the mind and one by Lord to the effect that rationality does not. I also consider a different, Kantian argument to the effect that rationality does not supervene on the mind.

ARTICLE HISTORY Received 5 August 2020; Accepted 5 August 2020

KEYWORDS responding to reasons; reason; quick objection; reasons first

1. Introduction

Philosophers often fail to respect the distinction between rationality and normativity. Here is just one example from a crucial moment in the history of modern moral philosophy. In arguing that there are no external reasons, Bernard Williams [1983: 110] says:

> There are of course many things that a speaker may say to one who is not disposed to φ when the speaker thinks that he should be, as that he is inconsiderate, or cruel, or selfish, or imprudent . . . But one who makes a great deal out of putting the criticism in the form of an external reason statement seems concerned to say that what is particularly wrong with the agent is that he is *irrational*. [original italics]

But a speaker who puts the criticism in the form of an external reason statement says that the agent has a reason to be disposed to φ. This is a normative statement. It does not impugn the agent's rationality at all.

CONTACT John Broome john.broome@philosophy.ox.ac.uk

The distinction between rationality and normativity has not been clearly drawn in philosophy. This paper begins to draw it. I first need to clarify the terms. The words 'rationality' and 'normativity' are caught in the complex tangle of words and meanings that derives from the Latin roots 'ratio' and 'norma'. They are surrounded by a lot of ambiguity. In order to say exactly what this paper is about, I first need to specify what I mean by these words. 'Rationality' and 'normativity' are the names of properties, which are ascribed to things using the adjectives 'rational' and 'normative'. I need to identify the two properties I mean to refer to. Doing so occupies the next three sections of this paper.

Having identified them, I shall go on to examine the substantive question of the relation between them. Since this paper is only a start, I shall concentrate on denying the extreme claim that the properties of rationality and normativity are the same. Though extreme, this claim is implicit in the widespread view that rationality consists in responding correctly to reasons. This view does not merely identify rationality with normativity; it implies that rationality is reducible to normativity. It is part of the 'reasons first' movement, which claims a foundational place for reasons in both normativity and rationality. It needs to be refuted.

2. The Meaning of 'Normative'

First, what do I mean by 'normative'? I mean: involving ought or a reason. I think this is the word's usual meaning in philosophy; it certainly is in moral philosophy. 'Normativity' in the title of this paper is in effect just standing in for 'ought' and 'reasons'. The subject of this paper is the relationship between rationality on the one hand and ought and reasons on the other.

Some philosophers give 'normative' the different meaning of involving correctness. What is the difference? Any rule and any intention sets up a standard of correctness. Complying with the rule or satisfying the intention is correct according to the rule or intention. So rules and intentions are normative in the sense of involving correctness. Furthermore, a person may be guided in her acts or thoughts by the standard of correctness. This is what happens when you follow a rule or fulfil an intention. For example, a man may follow the rule of wearing a tie on weekdays. Each weekday morning, guided by the rule, he may carefully and deliberately tie his tie. But it need not be the case that he ought to do this, or has a reason to do it. Nor need he even believe he ought to or has a reason to. He may have been brought up to do it, and never have given any thought to the rule. This rule is normative in the sense of involving correctness, but it is not normative in my sense.

Since 'normative' and 'normativity' are purely technical term, I am free to use them as I wish, so long as I define my meaning. I stick to their usual meaning in philosophy, but I find it convenient to extend this meaning a bit for the sake of making the term 'normative' more formally parallel to 'rational'. We do not usually apply the adjective 'normative' to a person, but I shall. I shall treat normativity as a property that a person may possess, like rationality. I shall say that a person is normative if she does whatever she ought to do, believes whatever she ought to believe, wants whatever she ought to want, and so on. That is to say: if she *F*s whenever she ought to *F*, where any verb phrase can be substituted for '*F*'. To continue the parallel, I shall also speak of 'requirements of normativity' in parallel with requirements of rationality: when a person ought to *F*, I shall say that normativity requires her to *F*.

My definition allows normativity to be a matter of degree. A person is normative to the extent that she *F*s whenever she ought to *F*. Degrees of normativity will be partially ordered. For example, if you are subject to the same requirements of normativity in two possible states of affairs, and in one of them you satisfy all the requirements you satisfy in the other and at least one more, then you are more normative in the first state than in the second. I would not expect degrees of normativity to be completely ordered.

Some philosophers are uncomfortable with my innovative use of 'normative'. It is a convenient piece of terminology, but I could have chosen another. For example, I could have chosen 'reason-respondingness' or better 'ought-respondingness' as alternative names for the property I refer to. The term is not important.

3. Meanings of 'Reason'

Next, what do I mean by 'rational'? Answering this question is a much longer task. Unlike 'normative', 'rational' is a term of common English. Philosophers are entitled to commandeer English words for technical purposes, just as the science of mechanics commandeered 'force'. However, doing so is risky because unwary readers may assume the word retains its common meaning. I therefore intend to use 'rational' with its common meaning. Moreover, I think this is what philosophers of rationality generally intend; they generally aim to analyse rationality as we ordinarily understand it.

The common meaning of 'rational' is complex. This section and the next are given over to clarifying it. The adjective 'rational' is cognate to the noun 'reason'. These words share the same Latin root 'ratio'. But this is not enough to tell us the meaning of 'rational', because 'reason' is very ambiguous word. We need first to identify the sense of 'reason' that 'rational' is cognate to. I shall approach this task historically.

The word 'reason' entered English from French along with the Norman invasion of England in 1066. After that invasion, nothing was written in English for more than a century. The version of English that subsequently arose is known as 'Middle English'. One of the first books written in Middle English was the *Ancrene Riwle*, whose earliest manuscript dates from about 1225. It contains all the earliest occurrences of the word 'reason' (spelled 'reisun') in English.[1] Even in the *Ancrene Riwle*, 'reason' was ambiguous. It appears with various different senses, all of which survive today.

Usually in that book, 'reason' means explanation. It still has this meaning today, for example in the sentence 'The reason the climate is changing is humanity's burning of fossil fuel'. An example from the *Ancrene Riwle* is:

> All strength comes from humility. And Solomon gives the reason why: where humility is, there is wisdom. [p. 125, folio 75]

Most often in the *Ancrene Riwle*, 'reason' refers to the special sort of explanation of a person's action that we nowadays call a 'motivating reason'. For example:

> This, now, is the reason of the joining: why Isaiah joins hope and silence, and couples both together. [p. 34, folio 19]

[1] One exception is an occurrence of 'reason' in *The Martyrdom of Sancte Katerine* from about the same date. There it has a sense that is now obsolete.

> Just once the *Ancrene Riwle* uses 'reason' to refer unequivocally to a normative reason, in this sentence:
>
> The third reason for fleeing the world is the gaining of heaven. [p. 73, folio 43].[2]

A closer representation of the original grammar is:

> The third reason of the world's flight is the gaining of heaven.

If the grammar is not enough to convince you that the author is referring to a normative reasons as opposed to a motivating one, the context should be. The author has previously promised to enumerate reasons why one ought to flee the world:

> Now hear reasons why one ought to flee the world: eight reasons at the least. [p. 72, folio 42]

A reason why one ought to *F* is a normative reason to *F*.

'Reason' in the *Ancrene Riwle* is a count noun. English also has a mass noun 'reason' with a normative meaning, but it did not appear till some centuries later. The *Oxford English Dictionary* (hereafter OED) lists no clear examples before this one from 1582:

> Yet there is reason to think, that they knew what they did as well as he. [Parsons 1852: 27]

Elsewhere in the *Ancrene Riwle*, 'reason' occurs once in the quite different sense that refers to a property of a person, specifically the faculty of reason. Since this is a mental property, I call this the 'mental sense' of 'reason' to contrast it with the normative sense. The original text is:

> Wummon is the reisun—thet is, wittes skile—hwen hit unstrengeth. [p. 121, folio 73]

This needs some exegesis. The author has just recounted a parable from the Bible. He is saying that the woman in the parable represents the faculty of reason. Perhaps because the word 'reason' had only recently acquired the mental sense, he glosses it using an older English term for the faculty of reason. The older term is: 'wittes skile', or 'wit's skill' in modern spelling.

Since this earliest use of 'reason' in the mental sense is obscure, here is a clearer one from Shakespeare:

> The will of man is by his reason sway'd.[3]9

Here is David Hume [1739–40: bk 2, part 3, sect 3] contradicting Shakespeare:

> Reason alone can never be a motive to any action of the will.

Another example from Hume is:

> 'Tis not contrary to reason to prefer the destruction of the world to the scratching of my finger. [ibid.]

This last remark has horrified some philosophers. It has been called 'silly' by Allen Wood [2013:65] and 'grotesque' by Michael Smith [2004: 76]. These authors understand Hume to mean (this is Wood's paraphrase):

> I can have no reason not to prefer the destruction of the world to the scratching of my finger.

[2] See Broome [2021] for an explanation of how this normative meaning came to be used.
[3] *A Midsummer Night's Dream*, act 2 scene 2.

But Hume did not mean that. He was using 'reason' in the mental sense, not in the normative sense that appears in Wood's paraphrase.

4. The Meaning of 'Rational'

So, from the birth of Middle English, 'reason' has had at least three senses: the explanatory sense, which includes motivating reasons as a prominent case, the normative sense and the mental sense. The adjective 'rational' appeared two centuries later. It is first recorded by the OED in 1398. From its beginning it was cognate to 'reason' in the mental sense, and in that sense only. At first it meant 'having the faculty of reason'. It had this meaning and no other for about two hundred years. The OED shows that for all that time it was applied as a predicate only of people, creatures, souls, minds and so on—all things that could possess the faculty of reason.

The noun 'rationality' appeared in 1628 as the name of the property that is ascribed by means of the adjective 'rational'. Since this property is just reason in the mental sense, 'rationality' and 'reason' in this sense were originally synonymous.

However, the meaning of 'rational' has subsequently broadened. 'Reason' in the mental sense refers only to a faculty. 'Rationality' today refers to the same faculty and also to the property of being in a particular state of mind—roughly, a state of mind that could have arisen from the exercise of the faculty of reason. The term 'structural rationality' is sometimes used by philosophers today for the rationality of states. These days, we would not count a person as fully rational if she had the faculty of reason but not structural rationality. For instance, a person is not fully rational if she does not intend means to her ends, even if she has the ability to ensure that she does intend means to her ends. Ability—faculty—is not enough for rationality; we expect results. Moreover, a person's rationality might improve even without her exercising an ability. She might come to intend an act that she believes is a necessary means to an end she intends, not by doing anything at all, but as a result of some subpersonal process within her. She is then more rational than she was before, even though she has not exercised her rational ability.

From 1598 onwards, the OED records 'rational' used as a predicate of things that do not have minds. These days we apply this predicate to acts, beliefs, city plans, and many other things without minds. These are derivative uses of 'rational'. They derive from the original, mental sense, which applies to people. The nature of the derivation varies with the object 'rational' is applied to, and it may be rather indefinite. For example, a person's act is rational if, were she to do it, she would be no less rational than if she were not to do it. A city plan is rational if it exhibits the sort of organization that is characteristic of a rational mind. And so on.

In its primary use, 'rationality' still denotes a property of a person. In this paper I stick to its primary use. In this use it retains one central feature even in its broadened sense that includes structural rationality. As a property of a person, rationality is specifically a mental property. Moreover, it depends on the person's other mental properties: rationality supervenes on the mind, as Ralph Wedgwood [2002] puts it. If a person would have the same mental properties apart from rationality in either of two possible situations, she would be equally rational in either.

For example, when a person intends to drink a glass of liquid, she is equally rational in the case when the liquid is petrol as she is in the case when it is gin, so long as the

difference is not registered in any mental property of hers.[4] For another example, if you fail to take a means to an end that you intend, this is not necessarily a failure of rationality if it is caused by some non-mental obstruction. It could be that, if you had had all the mental properties you do have, but the obstruction had not existed, you would have taken the means to your end. There would then have been no failure in your rationality, so supervenience implies there is no failure in the actual case either.

What about mental externalism? I believe the Taj Mahal is made of marble. Suppose that, elsewhere in the universe, there is a Twin Earth that has all the same intrinsic physical properties as Earth. On Twin Earth lives a person called 'John Broome'. His intrinsic physical properties, including the intrinsic properties of his brain, are the same as mine. He has a belief that he would express with the words 'The Taj Mahal is made of marble'. His belief is about the Twin Taj Mahal, whereas mine is about the actual Taj Mahal. His belief is therefore not the same as mine; it has a different content and beliefs are individuated by their contents. At least, that is the implication of externalism about mental content. I take each of a person's beliefs to be a mental property of hers. So Twin John's mind does not have all the same mental properties as mine has. If externalism is true, our mental properties do not supervene on our intrinsic physical properties, therefore.

For an analogy, think of the magnetic field of a particular magnet. The field belongs to the magnet, but other ferrous objects in the neighbourhood influence the field. The field's direction and strength at a particular point therefore do not supervene on intrinsic physical properties of the magnet. However, they are intrinsic properties of the magnet's field; indeed the field simply consists of the set of directions and strengths at all points. The analogy is this: a person corresponds to a magnet; the person's mind corresponds to the magnet's field; a belief corresponds to the direction and strength of the field at a point. A person's belief supervenes on the intrinsic properties of her mind, but it does not supervene on her intrinsic physical properties.

If externalism is correct, the principle that rationality supervenes on the mind does not imply that Twin John and I are exactly as rational as each other, since our minds have different properties. Nevertheless, we surely are exactly as rational as each other, so presumably there is some stronger principle that does have this implication. Presumably it would be a principle that rationality supervenes on internal properties of the mind, defined in some way or other. But I do not know any such principle, and I do not assert that one exists. At any rate, externalism is no threat to the principle I do assert, that a person's rationality supervenes on her mental properties apart from rationality itself. It simply suggests there is also a stronger principle.

4.1 Substantive Rationality

Some philosophers use 'rational' differently. They treat it as an adjective corresponding to 'reason' in the normative rather than the mental sense. For example, Niko Kolodny and John Brunero [2018] say:

[4] This famous example comes from Williams [1983].

'What would it be rational for an agent to do or intend?' could mean:

1. By doing or intending what would the agent make her responses (i.e., her attitudes and actions) cohere with one another? […]
2. What does the agent have reason, or ought she, to do or intend?

Definition 1 is supposed to pick out structural rationality. As it happens, the mention of actions prevents it from doing so, since a person's actions—apart from her mental actions—do not supervene on her mind. For example, suppose you intend to meet your obligations and believe that in order to do so you must return the book you borrowed from your friend. You intend to return the book. You put it in the mail to her, but the mail fails, so you do not return the book. You do not do as you intend; your actions do not cohere with your intentions. However, suppose you never find out about the failure, so your mind has all the properties it would have had had you returned the book. It may therefore be coherent. You may be structurally rational even though your action does not cohere with your intentions. So definition 1 is not accurate, but it is plainly meant as a definition of structural rationality.

Definition 2 also is inaccurate. Suppose you have reason to intend something but stronger reason not to intend it. Kolodny and Brunero do not mean to suggest that it would be rational in any sense for you to intend it. They must have meant to say 'conclusive reason'. I assume that is their meaning.

In any case, in definition 2 'reason' has its normative sense. The definition associates rationality with reason in this sense, so it gives 'rational' a normative meaning. This meaning could be etymologically justified. If you want an adjective cognate to 'reason' in the normative sense, 'rational' could serve. (An alternative is 'reasonable'.) However, so far as I can tell, this use of 'rational' is not historically justified. 'Rational' has never had this normative meaning in common English: it has always been cognate to 'reason' in the mental sense, and never in the normative sense.

It is not easy to be sure, since the extensions of the mental and normative senses coincide to a large degree. When an agent ought to do or intend something, it is usually rational in both senses for her to do or intend it. That is, she would usually be no less rational in both senses if she did or intended it than if she did not. This is because an agent normally has correct beliefs about what she ought to do or intend.

But we can separate the senses by taking a case where the agent ought to do or intend something, but believes she ought not to. Suppose there is nothing irrational about her false belief; it is supported by good—though misleading—evidence. Presented with a case like this, would ordinary English speakers use 'rational' in Kolodny and Brunero's second, normative sense? In this sense it would be rational for the agent to do or intend what she rationally believes she ought not to do or intend. Would any ordinary English speaker say that? Since mistaken normative beliefs are rare outside philosophers' examples, I cannot provide textual evidence. But I very much doubt that any ordinary English speaker would say it. Indeed, I would be surprised to find a philosopher saying it. I think it is a well-embedded conceptual feature of rationality, treated as a property of a person, that it is a mental property. I think this example makes that clear.

The normative sense of 'rational' is recent;[5] for many centuries at least, 'rational' had only the mental sense. So far as I can tell, the normative sense is an invention of philosophers. I mentioned in section 3 the risks of commandeering a common

[5] The term 'substantive rationality' apparently originated with Max Weber. See Kalberg [1980].

English word as a technical term. We do not need this sense of 'rational' in philosophy; we would do better to use a genuine technical term such as 'normative' in the way I suggested in section 2. The new sense of 'rational' simply leads to confusion. Most philosophers who write about rationality intend to write about it as it is commonly understood. That is my intention. Given all this, we should eschew the normative sense of 'rational', and I do.

I use 'rational' as cognate to 'reason' in the mental sense only, and I claim this is its correct usage when predicated of a person. But it does not matter whether or not I am right about this. Even if 'rational' may also be correctly used as a predicate of a person in a different sense, I use it in the specifically mental sense. This paper is about rationality as a mental property of a person.

4.2 Reified Rationality

But we do need to recognize a further meaning of 'rationality' in which it is not the name of a property at all. We sometimes reify the property of rationality, treating it as a thing rather than a property. We do the same for morality. Morality is the property a person possesses when she is moral, but sometimes we treat it as something that stands outside a person and makes prescriptions to her. Similarly we sometimes treat rationality as something that stands outside a person and makes prescriptions to her. The reified meaning of 'rationality' is well established in common English. It has its source in 'rationality's meaning as a property of a person, but it goes beyond that meaning.

In philosophy, reification is most apparent in the expression 'rationality requires'. The word 'requires' has different functions. It may be used to specify a necessary condition for possessing a property. That is its function in 'Cleanliness requires soap': having soap is a necessary condition for being clean. But in 'The law requires you to vote' 'requires' specifies a prescription: the law prescribes that you vote. 'Rationality requires you to intend means to your end' might be interpreted in either of these ways. It might mean simply that intending means to your end is a necessary condition for your having the property of rationality. With this meaning, rationality is not reified. Or it might mean that rationality prescribes that you intend means to your end. With this interpretation, rationality is reified. It is treated as a thing like law.

I claim that the expression 'rationality requires' is most naturally understood in the reified sense.[6] I think we would not naturally say that rationality requires you to be alive. Being alive is a necessary condition for possessing the property of rationality, so unreified rationality does indeed require you to be alive. On the other hand, rationality does not prescribe that you are alive, so reified rationality does not require you to be alive. We naturally take 'rationality requires' this second way. What rationality requires in this more natural sense is a subset of the necessary conditions for being rational. I call this the set of 'requirements of rationality' or 'rational requirements'.

Nevertheless, satisfying requirements of rationality is sufficient for possessing the property of rationality: if you satisfy all the requirements of rationality you are fully rational. Moreover, the degree to which you satisfy requirements of rationality is the degree to which you have the property of rationality. These degrees are partially ordered. For example, if in one possible situation you satisfy all the requirements of

[6] My evidence is contained in Broome [2013: 119–26]. Here I give just one example.

rationality that you satisfy in another, and you also satisfy at least one more, you are more rational in the first situation than in the second.

5. The Identity Claim

I have described what I mean by 'normativity' and by 'rationality'. Now I come to the substantive question of the connection between these properties. Most of the rest of this paper is devoted to the popular view that is expressed by the formula 'rationality consists in responding correctly to reasons'. A part of this view is that the property of rationality is identical to the property of responding correctly to reasons. But the words 'consists in' go further than 'is'; they claim more than identity. They imply that rationality can be reduced to reasons, which they imply are metaphysically more fundamental. The formula that rationality consists in responding correctly to reasons is part of the 'reasons first' movement, which tries to reduce rationality and all of normativity to reasons. I aim to refute the reductive claim by means of refuting the identity claim, which is weaker.

What is the property of responding correctly to reasons? I take it to be doing whatever your reasons require you to do, believing whatever your reasons require you to believe, and so on—in short, *F*ing whenever your reasons require you to *F*. This is not the same as *F*ing whenever you have a reason to *F*. When you have a reason to *F*, you may have a stronger reason not to *F*, and in that case *F*ing is plainly not responding correctly to reasons. When your reasons require you to *F*, another way of putting it is that you ought to *F*. So responding correctly to reasons is *F*ing whenever you ought to *F*. This is just the property of normativity as I defined it in section 2. So the claim I aim to refute is that the property of rationality is the same as the property of normativity. I shall sometimes call it 'the property-identity claim' and sometimes just 'the identity claim'.

Responding correctly to reasons might be understood as a narrower property than *F*ing whenever you ought to *F*. It might be understood as *F*ing whenever you ought to *F* because of the reasons that make it the case that you ought to *F*. Adopting this narrower interpretation would not affect the conclusions of this paper, but would add to its complexity. So I have decided to stick to the simpler interpretation.

The property-identity claim is implied by the different claim that what rationality requires of you is necessarily the same as what normativity requires of you, which is to say what you ought. I call this 'the requirement-identity claim'. To check this implication, suppose the requirement-identity claim is true. Suppose also that you are rational. Then you satisfy all the requirements of rationality you are under. Since these requirements are the same as the requirements of normativity, you satisfy all the requirements of normativity you are under, which means you are normative. So, necessarily, if you are rational you are normative. The same argument in reverse shows that, necessarily, if you are normative you are rational. The properties of being rational and being normative necessarily have the same extension. According to a standard criterion for identity of properties, this means they are the same property (see See Orilia and Swoyer [2020]).

Normativity and rationality and both properties that have degrees. Does the requirement-identity claim imply that, when you are less than fully rational and less than fully normative, your degree of rationality is the same as your degree of normativity? To put this more accurately: does the requirement-identity claim imply that the

partial ordering of degrees of normativity exactly matches the partial ordering of degrees of rationality? It does not. To derive this conclusion from the requirement-identity claim, we would have to make some further assumptions. Though these assumptions would be very plausible, I shall not make them because my argument in this paper does not touch on degrees of normativity or rationality.

The property-identity claim as I mean it is the claim that the property of being fully normative and the property of being fully rational are the same. This is implied by the requirement-identity claim. I shall argue that the property-identity claim is false. This implies that the requirement-identity claim is also false.

My argument against the property-identity claim is that rationality supervenes on the mind but normativity does not. If your mental properties (apart from rationality) are the same in one possible state as they are in another, you are equally rational in the two states, but you may not be equally normative. Rationality and normativity therefore cannot be the same. I call this 'the quick objection' to the identity claim (see Broome [2013: ch. 5]).

There are only two ways to oppose the quick objection. One is by asserting that normativity supervenes on the mind. The other is by denying that rationality supervenes on the mind. Neither of these responses is sufficient to prove the identity claim, but refuting both is sufficient to refute the identity claim. The next two sections aim to refute both in turn.

6. Normativity Does Not Supervene on the Mind

Normativity is the property of *F*ing whenever you ought to *F*. It will supervene on your mind if two conditions are true. The first is that what normativity requires of you—what you ought—supervenes on your mind. To be clear: by this I mean that, for any *F*, whether or not you ought to *F* supervenes on your mind. The second is that, whenever you ought to *F*, your *F*ing supervenes on your mind. To be clear: whether or not you *F* supervenes on your mind. These are not strictly necessary conditions for normativity to supervene on your mind, but if they do not obtain, normativity could supervene on your mind only by good luck.

I shall examine each condition in turn.

6.1 Does Whether or Not You Ought to F *Supervene on your Mind?*

The claim that what you ought supervenes on your mind is a sort of subjectivism about ought. It can be supported by any of several openly subjectivist theories. For example, one is the theory that you ought to *F* if and only if *F*ing has the greatest expected value for you out of all the alternatives, where expected values are given by your own credences and your own judgements of value.

Another theory starts from subjectivism about reasons. It claims that all your reasons are states of your mind. For example, reasons might be pairs, each consisting of a desire and a belief. Add to this the claim that your reasons determine what you ought by weighing against each other on the basis of your subjective judgements of weight. The result is a subjectivist theory of ought.

Many philosophers find subjectivism about reasons unattractive. It conflicts with common sense if nothing else. Common sense tells us that facts about the external world can be reasons. For example, the fact that heavy clouds are gathering is a

reason to expect rain. Benjamin Kiesewetter agrees with common sense in this respect. He thinks that many reasons are facts about the external world. Here is one of his examples. Suppose you are hunting a murderer, and you see someone disappear behind a tree. That someone disappeared behind the tree, which is a fact about the external world, may be a reason for you to believe that the murderer is hiding behind the tree [Kiesewetter 2017: 167 ff.].

Nevertheless, Kiesewetter [ibid.: ch. 7] argues that what you ought supervenes on your mind. So he claims that this sort of subjectivism about ought is consistent with the view that reasons are often facts about the external world.

To support this claim, he says, first, that what you ought depends only on those of your reasons that are available to you. A reason is available to you only if it is part of your body of evidence. You may have reasons that are not available to you, but we can ignore those ones because—Kiesewetter assumes—available reasons are the only ones that contribute to determining what your ought [ibid.: ch. 8].[7] From now on in discussing Kiesewetter, I use 'reason' to refer to available reasons only.

The fact that someone disappeared behind the tree is an available reason for you only if it impinges on your mind to the extent of being part of your evidence. This makes it possible for the existence of an available reason to supervene on your mind even though the reason is external.

But it only opens up a possibility; it does not ensure that the existence of an available reason supervenes on your mind. Change the example. Suppose now that you do not actually see someone disappearing behind a tree. It seems to you that you do, but this appearance is illusory. Actually, no one disappears behind the tree. Then you do not have the reason I described for believing the murderer is behind the tree. That reason was the fact that someone disappeared behind the tree, but in the new case there is no such fact. Yet your mind is in exactly the same state in the two versions of the example. So the existence of the reason does not supervene on your mind.

Kiesewetter offers two alternative responses to this problem. The first is a strong sort of externalism about the mind. In the original version of the example, the fact that someone disappeared behind the tree is part of your evidence; in the second version it is not. According to Timothy Williamson's [2000: ch. 1] externalist theory, your evidence is what you know, and your knowledge is a mental state of yours. So you are not in the same mental state in the two cases. Yet I assumed you are; that is how I demonstrated that the existence of your reason does not supervene on your mind. If Williamson is right, my demonstration fails.

So Kiesewetter's view that a reason must be part of your evidence, together with Williamson's externalism about the mind, may be enough to ensure that the existence of a reason supervenes on your mind. They are not yet enough to ensure that what you ought supervenes on your mind. What you ought depends also on what your reasons are reasons for, and how they combine together.

But we might extend Kiesewetter's theory to the extent of claiming that what you ought depends only on your total body of evidence, perhaps together with other features of your mind. Your evidence will include external facts that constitute reasons. If we now add Williamson's externalism about the mind, so that your body of evidence is

[7] I take it that a fact is not a reason at all unless it contributes to determining what you ought—see the definition of a reason in Broome [2013: ch. 4].

part of your mind, we shall get the conclusion that what you ought supervenes on your mind.

In this way, Kiesewetter combines a sort of subjectivism about ought with reasons that are facts about the external world. It strikes me as a sort of sleight of hand. The argument extends your mind to include the external facts that are reasons. The fact that someone disappeared behind the tree, when it is a fact, is a feature of your mind. But Williamson's version of externalism about the mind is unappealing. It is not externalism about the content of mental states, which I mentioned in section 4 and which is widely accepted. It is externalism about the existence of a mental state: your knowledge that someone disappeared behind the tree is a mental state of yours, and whether or not you have this mental state depends on whether or not someone disappeared behind the tree. Like subjectivism about reasons, this sort of externalism about the mind offends common sense.

Kiesewetter recognizes that many people do not accept it. So he offers an alternative argument. Go back to the example. In the first version you have a reason to believe the murderer is behind the tree, which is the fact that someone disappeared behind the tree. In the second version, you do not have that reason to believe the murderer is behind the tree, but Kiesewetter claims you do have a different reason to believe it, namely that it seems to you that someone disappeared behind the tree. Kiesewetter claims this second reason is just as strong as the first. He offers us this 'backup view':

> If *A*'s total phenomenal state supports *p*, and *p* would—if true—be an available reason for (or against) believing *q*, then *A*'s appearances provide an equally strong available reason for (or against) believing *q*. [Kiesewetter 2017: 173]

If this is true, the reason you have in the second version of the example (the appearance) is just as strong a reason to believe the murderer is hiding behind the tree as the reason you have in the first version (the fact that someone disappeared behind the tree). Therefore, what your reasons require—what you ought—is the same in the two cases. So even if we drop externalism about the mind and accept that your mind is the same in the two cases, the example is consistent with the supervenience of what you ought on the mind.

The backup view faces at least two difficulties. One is that you have the appearance in the first version of the example as well as in the second. If it is a reason in the second version, it must be just as strong a reason in the first. But in the first version you have a further, external reason, which is the fact. If, as Kiesewetter supposes, your reasons in the second version are just as strong as in the first, this seems to imply that the fact has no weight as a reason. Only the appearance counts for anything, and the external reason counts for nothing. This makes Kiesewetter's appeal to external reasons seems like a sham.

Second, an amendment to the example shows that the backup view is false if we accept, as Kiesewetter does, that the external fact can be a reason. Let us add two assumptions to the example: first, you know that no one is nearby apart from you and perhaps the murderer; second, you have received a fairly reliable report that the murderer was recently seen in a distant city. In the first version of the example, it is a fact that someone disappeared behind the tree, and this fact is supposed to be a reason to believe the murderer is hiding behind the tree. It is nothing less than a conclusive reason. You know that someone disappeared behind the tree and the murderer is the only person it could be, so you definitely ought to believe the murderer is hiding

behind the tree. But in the second version, it only appears to you that someone disappeared behind the tree. This appearance is supposed to be a reason to believe the murderer is hiding behind the tree, but you do not have a conclusive reason to believe this. Given the report that the murderer is elsewhere, it is not the case that you ought to believe the murderer is hiding behind the tree. So if both the fact and the appearance can be reasons, the appearance is definitely not as strong a reason as the fact. The conclusion is that, in the first version of the example, you ought to believe the murderer is hiding behind the tree whereas in the second version that is not so. Yet unless we accept externalism, your mind is the same in both cases. So once again, what you ought does not supervene on your mind.

Kiesewetter claims that what you ought supervenes on your mind, even though some reasons are facts in the external world. His argument from the backup view fails. I conclude that his defence of this claim has to depend on an unappealing sort of externalism about the mind.

This does not mean that I comprehensively reject subjectivism about ought, because there are other subjectivist theories besides Kiesewetter's. At the beginning of this section I mentioned one that does not refer to reasons. I do not have arguments against all of them. So my argument that normativity does not supervene on the mind rests mainly on what is coming next.

6.2 Your Fing Does Not Supervene on Your Mind Whenever You Ought to F.

Acting on the world outside your mind does not supervene on your mind. For example, raising your arm does not supervene on your mind. You might fail to raise your arm even while your mind has exactly the properties it would have if you raised it. Your nerves might fail to activate your muscles and you might be looking the other way. So if you ever ought to act on the outside world, it will not be true that, whenever you ought to *F*, *F*ing supervenes on your mind.

But often you ought to act on the outside world. For example, it can be the case on some occasion that you ought to insure your house against fire. Suppose you ought and you believe you ought. Suppose you set about insuring your house. You complete an application form and pay a premium to an insurance company in the usual way, without having studied all the fine print carefully. Now take two different cases. In the first, everything proceeds as expected, and your house is insured. In the second case the small print contains a clause that says your house is insured only if its roof is constructed of slate, tiles or metal. Actually your house's roof is constructed of cedar shingles, so the house is not insured. Suppose this fact never comes to your attention because there is no fire. Then your mental properties are exactly the same in both cases. Yet in one case you insure your house as you ought and in the other you do not. You may be normative in one but not in the other. So your normativity does not supervene on your mind.

The only way to argue that, whenever you ought to *F*, your *F*ing supervenes on your mind is to deny that it is ever true that you ought to act on the external world. We would have to deny that you ought to insure your house, for instance. We might say instead that you ought to bring yourself to believe you have insured your house. Or it might be that you ought to intend to insure your house, or to have some other mental property.

But this is utterly implausible, in the example and in general. It could easily be true that you ought to insure your house. You ought to make sure your car's brakes are in good condition, you ought to be kind to strangers, look both ways before you cross the road, and so on—all acts in the external world. We accept innumerable ordinary normative claims like these. It is quite implausible that all of them are false.[8]

To summarize this section: the claim that normativity supervenes on your mind depends on two conditions. The first can be defended only on the basis of unappealing philosophical theories. The second is utterly implausible. So we may safely conclude that normativity does not supervene on the mind.

7. Rationality Supervenes on the Mind

I next respond to two very different ways of arguing that rationality does not supervene on the mind.

7.2 Lord's Argument

Errol Lord [2017] favours the property-identity claim. He adopts the first way of opposing the quick objection; he denies that rationality supervenes on the mind. He thinks that rationality consists in responding correctly to the reasons you possess, that you possess a reason just when you are in a position to know it (see also Lord [2010]), and that being in a position to know a reason, as he understands it, does not supervene on your mind.

Lord uses the example of drinks to illustrate possessing a reason. You intend to drink a glass of liquid, which is either petrol or gin. Lord adds the assumption that, lying on the counter right in front of you, is an authoritative card that specifies which the liquid is. If it is petrol, the fact that it is petrol is a reason not to intend to drink it. The fact that the reason is described on a card right in front of you—and Lord adds some further supporting details—implies that this reason is possessed by you.

However, Lord supposes you do not read the card so that the reason does not impinge on your mind, even though it is possessed by you. Let us suppose the glass of liquid is taken away before you put your intention to drink into effect, so you never find out what is in it. Then your mind has exactly the same properties at all times whether the liquid is petrol or gin, even though in the former case you possess a reason not to intend to drink the liquid, whereas in the latter case you do not.

Lord assumes that, if the liquid is petrol, the reason you possess not to intend to drink it outweighs any opposing reasons you might possess, so responding correctly to the reasons you possess entails not intending to drink the liquid. Since you intend to drink it you are therefore irrational according to his theory of rationality.

[8] Kurt Sylvan has pointed out to me that this is precisely what H. A. Prichard [2002] claims. Even with Jonathan Dancy's help I have not been able to extract a credible argument from Prichard's text. Jesse Hambly has pointed out to me that T. M. Scanlon [1998: 21] says 'Judgement sensitive attitudes constitute the class of things for which reasons in the standard normative sense can sensibly be asked or offered'. This seems to deny that there can be reasons for acting in the external world. However, Scanlon immediately goes on to discuss reasons for action, without denying there are such things. Furthermore, he discusses reasons for action through the rest of the book. I think he means to that a reason for *F*ing, where *F*ing is an action, is also a reason for intending to *F*.

Why should we accept that? Lord evidently thinks he has supplied enough details to convince us that your failure to read the card is a failure of rationality. For the sake of argument, let us accept it.

It yields the conclusion that rationality does not supervene on the mind only if you would not be equally irrational were the liquid gin and you intended to drink it. But in that case too, you would fail the read the card in front of you. So if your failure is irrational in one case it is irrational in the other. Lord has not presented us with two cases where you are irrational in one and not irrational in the other.

What if the details are such that, in the petrol case, you know something you do not know in the gin case? For example, suppose you know the card contains information that is vital for your health. In the gin case, you do not know this because it is not true. Could this mean you are irrational in one case and not the other, even though your mind is the same in both cases?

No. If Williamson [2000: ch. 1] is right that knowledge is a mental state, you have different mental properties in the two cases; your mind is not the same in the two cases. If Williamson is wrong, your mind may indeed have the same properties in the two cases. If it does, then in the gin case you believe the card contains information that is vital for your health, just as you do in the petrol case. Given that your failure to read the card is irrational in the petrol case, it is irrational in the gin case too.

Despite what he intends, Lord's example only makes it clearer that rationality supervenes on the mind. As I said in section 4, it is a well-embedded conceptual feature of rationality as we ordinarily understand it that it is a mental property.

7.2 A Kantian Argument

There is a quite different way of denying that rationality supervenes on the mind. It is Kantian in nature, but I do not claim that it represents Kant's own view. I shall approach it indirectly, by a detour through a puzzle about morality.

A Kantian view is that morality supervenes on the mind. The view is that to be moral you require only a good mind—specifically, a good will. If, by bad luck, your good will does not achieve good consequences for the world, you are no less moral for that.

Even if morality supervenes on the mind, it does not follow that normativity does. Morality is only a part of normativity. My examples were about prudence rather than morality, and there is no suggestion that prudence supervenes on the mind. So the conclusions I drew from the examples are not affected. The point I want to make is different. Applied to morality, the Kantian view seems plausible, and it raises a puzzle.

It seems plausible that your possession of the property of morality does indeed supervene on your mind. If you intend to act well but bad luck intervenes and prevents you from achieving a good result, it seems plausible that your failure does not count against your morality. You cannot be blamed, so you cannot be any less moral. In general, you cannot be less moral in one state than in another if your mind is no different. Yet on the other hand, it is implausible that everything morality requires of you—in other words, everything you morally ought—supervenes on your mind. For example, morality requires you not to cause unnecessary suffering, and whether or not you cause unnecessary suffering does not supervene on your mind. Morality is surely aimed at the world, not at improving your own mind. So it seems that

morality supervenes on the mind but that what morality requires does not. How can these claims be reconciled? That is the puzzle.

A solution is to recognize that the reification of morality in our conceptual scheme has progressed a long way. In the first instance, 'morality' is the name of a property that people possess. Plausibly, this property supervenes on the mind. If we construe 'morality requires' on the model of 'cleanliness requires', what is morally required of you is whatever is a necessary condition for possessing the property of morality, and nothing more. If morality supervenes on the mind, what is morally required in this sense also supervenes on the mind.

But actually we reify morality and treat it as an external entity that makes prescriptions to us. I shall use the capitalized word 'Morality' for the name of this reified entity. Then 'Morality requires' should be construed on the model of 'the law requires' rather than 'cleanliness requires'. It means much the same as 'Morality prescribes'. Reified Morality does not require of you everything that is a necessary condition for possessing the property of morality. For example, a necessary condition for being moral is to be alive, but Morality does not require you to be alive. Morality does not prescribe being alive.

We would not naturally say that morality requires you to be alive. This shows that in the expression 'morality requires', 'morality' most naturally refers to reified Morality rather than unreified morality.

One consequence of reification is the one I have just described, that Morality does not require everything that is a necessary condition for being moral. But our reification of Morality goes further than that. Reified Morality also requires some things of you that are not necessary conditions for being moral. It requires some acts and omissions in the external world. For instance, Morality requires you to make sure your car's brakes are in good condition, and it requires you to refrain from murder. This is part of our common-sense understanding of Morality, and Kantians would not deny it. Being kind to strangers and refraining from murder do not supervene on your mind. Since we are making the Kantian assumption that the property of morality supervenes on the mind, it follows that they are not part of the property of morality. So on the Kantian view, reified Morality in some ways goes beyond morality as a personal property. This solves the puzzle: morality supervenes on the mind but what Morality requires does not.

Morality is not particularly a subject for this paper. Its relevance is that it could provide a model for the reification of rationality. We do indeed reify rationality, as I said in section 2. Let capitalized 'Rationality' be the name of the reified entity. Another name for it is 'Reason'. If Rationality is reified to the same degree as Morality, the requirements of Rationality need not supervene on the mind any more than the requirements of Morality do, even though the property of rationality does supervene on the mind.

This makes it possible for Rationality to require acts or omissions in the external world. To take one example, Rationality might require you to act only in accordance with that maxim through which you can at the same time will that it become a universal law [Kant 1785: 88]. I earlier argued that rationality cannot be identical to normativity because rationality supervenes on the mind whereas normativity does not. But the argument does not apply to reified Rationality if it does not supervene on the mind. It could even turn out that Rationality and normativity are identical. It could turn out that what you ought to do is nothing other than what Rationality requires of you.

There would be a lot more work to do before this Kantian claim could be established. First, it needs to be explained why the requirements of Rationality do not supervene on the mind. It is plausible that the requirements of Morality do not, because Morality is concerned with the world rather than your mind. The same cannot be said of Rationality, however reified, because it is much more concerned with your mind. It is strange to claim that Rationality may require something that is not a necessary condition for being rational. Second, it needs to be explained why Rationality encompasses the whole of normativity. Even if it is granted that Rationality does not supervene on the mind, that is far short of the conclusion that it constitutes all of normativity.

Those are two big jobs to do. But if they could be done, it would mean there is a concept of Rationality in which it is identical to a concept of normativity.

But this concept of Rationality would be very distant from rationality, the property that is possessed by people. Take a new example. Suppose Rationality requires you to act only in accordance with that maxim through which you can at the same time will that it become a universal law. I take this to mean that Rationality requires of you that you do not perform any act unless it conforms to some maxim that you can will to become a universal law. This requirement does not supervene on your mind. For example, suppose you give money to an honest-looking person who is soliciting contributions on the street. You believe this act conforms to the maxim 'Give money to charity', which you can will to become a universal law. Now distinguish two cases. In the first, the person is indeed collecting for charity, so your act conforms to your maxim and does not violate the requirement of Rationality. In the second case, the person is collecting for a terrorist organization, and giving money to a terrorist organization does not conform to any maxim that you can will to become a universal law. In this second case you do violate the requirement of Rationality. Yet your mind might have all the same properties in both cases. You might never find out that in the second case you contribute to a terrorist organization; you might go to your grave believing you gave to charity. You are plainly equally rational in the two cases because your rationality supervenes on your mind. The cases differ in what reified Rationality requires of you, but your unreified rationality is the same in both.

So the Kantian project of reification does nothing to bring unreified rationality closer to normativity. It remains a mistake to identify these very different properties.

8. Conclusion

I conclude that the property-identity claim is false. It is false that rationality is responding correctly to reasons. *A fortiori*, it is false that rationality consists in responding correctly to reasons. This reductive claim fails.

The Kantian argument in section 7 does not aim to support a reductive claim about rationality. Indeed, it tends in the opposite direction towards supporting the view that normativity is reducible to a sort of reified Rationality. Since this sort of reified Rationality is far removed from the property of rationality, the failure of the property-identity claim tells us little about this Kantian reduction.

Kiesewetter and Lord support the claim that rationality consists in responding correctly to reasons. There are other more or less similar reductive claims about rationality. For example, Derek Parfit [2011: 111] says that rationality consists in responding correctly, not just to reasons, but also to apparent reasons. Moreover, Parfit does not

interpret responding correctly to reasons as *F*ing whenever your reasons require you to *F*, but as *F*ing or trying to *F* whenever your reasons require you to *F* [ibid.: 22]. Still, his view is an attempt to reduce the property of rationality to reasons. I think this reductive enterprise is mistaken, but my argument in this paper is not a comprehensive answer to it. In this paper, I have argued only against the identity claim itself. My argument has been only a 'quick objection' as I call it.[9] It succeeds against the identity claim, but other reductive claims call for less quick objections, which I have presented in my book *Rationality Through Reasoning* [Broome 2013: chs 5, 6].

I recognize there are some tight connections between rationality and normativity. For one thing, rationality may be a source of normativity: if rationality requires you to *F*, that may be a reason for you to *F*. Furthermore, rationality requires you to intend to *F* whenever you believe you ought to *F*. I call this requirement of rationality 'enkrasia' [ibid.: section 9.5]. It requires you to respond to your normative beliefs in a particular way. It differs from the identity claim because it is only one of many requirements of rationality; another is the requirement to intend whatever you believe is a means implied by an end you intend, and there are many others. Still, it does constitute a tight connection between rationality and normativity.

Acknowledgments

My thanks to Ruth Chang, Olav Gjelsvik, Benjamin Kiesewetter, Philip Pettit, Franziska Poprawe, Wlodek Rabinowicz, Kurt Sylvan and particularly Jesse Hambly and Peter Railton for very helpful comments. Research for this paper was supported by ARC Discovery Grants DP140102468 and DP180100355.

Disclosure Statement

No potential conflict of interest was reported by the author.

References

Ancrene Riwle, English Text, http://www.bsswebsite.me.uk/History/AncreneRiwle/AncreneRiwle.html.

Broome, John 2013. *Rationality Through Reasoning*, Wiley–Blackwell.

Broome, John, Forthcoming. The First Normative 'Reason', in *Normativity, Rationality and Reasoning: Selected Essays*, Oxford: Oxford University Press.

Day, Mabel, ed., 1952. *The English Text of the* Ancrene Riwle, edited from Cotton MS Nero A XIV, Oxford: Oxford University Press.

Hume, David 1739–40 (1888). *A Treatise of Human Nature*, ed. L. A. Selby-Bigge, Oxford: Oxford University Press.

Kalberg, Stephen 1980. Max Weber's Types of Rationality: Cornerstones for the Analysis of Rationalization Processes in History, *American Journal of Sociology* 85: 1145–79.

Kant, Immanuel 1785 (1948). *Groundwork of the Metaphysics of Morals*, translated as *The Moral Law* by H. J. Paton, Hutchinson's University Library.

Kiesewetter, Benjamin 2017. *The Normativity of Rationality*, Oxford: Oxford University Press.

Kolodny, Niko, and John Brunero 2018. Instrumental Rationality, in *Stanford Encyclopedia of Philosophy*, ed. Edward Zalta, https://plato.stanford.edu/entries/rationality-instrumental/.

Lord, Errol 2010. Having Rreasons and the Factoring Account, *Philosophical Studies* 149: 283–96.

Lord, Errol 2017. What You're Rationally Required to Do and What You Ought To Do (Are the Same Thing!), *Mind* 126. 1109–54.

[9] Originally in Broome [2013: section 5.2].

Orilia, Francesco, and Chris Swoyer. 2020. Properties, in *Stanford Encyclopedia of Philosophy* (Summer 2020 Edition), ed. Edward Zalta, https://plato.stanford.edu/archives/sum2020/entries/properties/.

Oxford English Dictionary Online, http://www.oed.com

Parfit, Derek 2011. *On What Matters*, volume 1, Oxford: Oxford University Press.

Parsons, Robert 1582. *A Defence of the Censure Gyuen vpon Two Bookes of W. Charke and M. Hanmer Against E. Campian*.

Prichard, H. A. 2002. Duty and Ignorance of Fact, in his *Moral Writings*, ed. Jim MacAdam, Oxford: Oxford University Press: 84–101.

Scanlon, T. M. 1998. *What We Owe to Each Other*, Cambridge MA: Harvard University Press.

Smith, Michael 2004. Humean Rationality, in *The Oxford Handbook of Rationality*, eds Alfred R. Mele and Piers Rawling, Oxford: Oxford University Press: 75–92.

Wedgwood, Ralph 2002. Internalism Explained, *Philosophy and Phenomenological Research* 65: 349–69.

Williams, Bernard 1983. Internal and External Reasons, in his *Moral Luck*, Cambridge UK: Cambridge University Press: 101–13.

Williamson, Timothy 2000. *Knowledge and its Limits*, Oxford: Oxford University Press.

Wood, Allen W. 2013. Kant on Practical Reason, in *Kant on Practical Justification: Interpretive Essays*, eds Mark Timmons and Sorin Baiasu, Oxford: Oxford University Press: 57–86.

Silke Schicktanz

Normativity and Culture in the Context of Modern Medicine: A Prospective Vision of an Elective Affinity

> Cursed, cursed creator! Why did I live? Why, in that instant, did I not extinguish the spark of existence which you had so wantonly bestowed?
> – The 'monster' in Mary Shelley's *Frankenstein*; ch. 15

Mary Shelley's romantic novel *Frankenstein* (1818) symbolizes a particular position towards the means and achievements of modern life sciences: It is marked by hopes and fears related to the outcomes of such research as well as an ambivalent admiration for the courage of those inventors' desire to try out what is at least thinkable. Apart from the rich ambivalence of this novel, it is a kind of irony that 'Frankenstein' as term still appears in many public debates on genetic modification, transplantation medicine, or synthetic biology as a label to classify such innovations as repugnant and dismissible. Hereby, the creator and the creation are almost equated.

Almost two centuries later, pop-cultural phenomena such as the 2011 Hollywood movie *Rise of the Planet of the Apes*, directed by Rupert Watt, show, in a similar vein, the subtle ways in which science fiction serves as platform to address various ethical issues and concerns. This film tells the story of the application of a new somatic gene therapy targeting dementia in a near future. When applied in human primates it significantly boosts the primates' social and cognitive intelligence. Feeling inspired by the outstanding intelligence performances of the tested chimpanzees, the young leading researcher (here rather a failed hero than a mad scientist) is tempted to apply the new drug to his demented father. He does so by bypassing the common professional ethics of informed consent of his father or any review board assessment, though his actions seem justified as a case of 'ultimate ratio,' and by his passionate love for his father. While the movie overall presents more fiction than science, it touches upon many ethical concerns related to modern science: From animal testing and the underlying arbitrariness of a human-ape distinction, to the ethics of inserting artificially genes into the genome of a species, and finally experimental drug testing, the movie seems to leave nothing out.

Considering both, *Frankenstein* and *Rise of the Planet of the Apes*, what is most relevant is the observation that novels and other types of cultural narratives

https://doi.org/10.1515/9783110669398-016

allow the expression of concerns, feelings, or worldviews otherwise easily neglected by the rational language of scientific enthusiasm. These examples mark a period of almost two hundred years of cultural reflection on science and medicine, each embedded within a framework of understanding of science and narrative techniques typical to its own era.

It is exactly this interplay between the socio-cultural dimension and the norms expressed that is worth reconsideration. In the following, I use the terms 'socio-cultural' and 'cultural' synonymously as umbrella terms, to signify 'culture' in a wide sense. By this, despite the difficulties in defining such a highly loaded term, I mean meanings, understandings, and practices 'presenting' or 'symbolizing' our interactions with other beings or the environment by verbal or non-verbal forms. According to my definition, 'culture' stands here for the robust and historically (partly) stable system of ideas and practices within a collective, expressed in various forms of mediality.[1] By this definition, culture differs significantly from psychological attempts or spontaneous social inter-individual interactions.

Culture overlaps with economics, politics, and law, which are also collective, public endeavors, but is less explicit than those orders and rules. It is – to use Clifford Geertz's metaphor – a "web" (1973) in which we collectively feel cocooned without often being aware of it. Culture exists only when we give it a cultural meaning – and we tend to do so, when others ask for explanations of things that are not self-evident. However, I am aware – and it might be particular methodologically relevant – that social interactions and cultural forms can strongly differ with regard to their material, spatial, and temporal validity. The study of culture – as I here understand it – can therefore range from ethnographic, anthropological, sociological, or empirical-ethical approaches to the study of literature, media, or art.

For instance, a cultural study of German science fiction novels of twentieth century and an empirical study of social interactions by non-participatory observation of doctor-patient-communication during cancer care differ significantly with regard to their methodological accounts and, perhaps, theoretical assumptions. However, what these studies share (or can share, according to the idea I defend here) is that they allow us insights in the often-hidden meta-structure of providing 'meaning' to the way how norms or values and medical practices or ideas regarding the human body are mutually shaped in a particular setting (see Schicktanz 2007). The reservoir of cultural attempts can help us to reflect upon

1 In a broad sense, I am interested in the parallels of different media (including literature, movies, artistic performances, or even images and paintings), but of course I am aware that each media has its historical meanings and methodological constraints.

both, the hidden as well as the explicit meanings that the body 'has' and that inform so even our normative reflections of what one ought to do with the body.

The relationship between applied ethics and cultural studies is not an easy one to determine. The juxtaposition of culture and normativity can provoke objection, when norms and values are seen already as part of 'culture.' However, such an almighty concept of culture seems almost impenetrable and therefore unproductive for any reflective study. Still, there exist many ways to conceptualize the relationship between culture and normativity. At one end of this spectrum, as in post-modern cultural studies, there is a trend to see things from a social-constructivist point of view (Clifford and Marcus 1986; Bauman 1993). In such a view, all explicit and hidden values and norms are construed as arbitrary, local, or contingent. Any underlying moral message is seen as just one of many possible messages, and the body is understood as a text that can be rewritten and reinterpreted in various terms. At the other end of the spectrum, where analytical (Searle 1995) or structuralist approaches (Douglas 1970, 1992; Lévi-Strauss 1961, 1963) are located, the cultural narrative provides a vehicle to infect its listener with 'true' propositions of what is and what ought to be. However, whether one of these extreme positions really does justice to the normativity embedded and entangled with cultural interpretations of our bodily practices and images must be critically questioned.

In the following, I suggest a third, alternative way to conceptualize and analyze the productive joints and links between current approaches of applied ethics and socio-cultural studies. The aim of the approach I propose is to open up future cultural studies for an 'ethical turn,' but not in the naïve sense of ethics, which conflates it with a pre-fixed set of norms and values (whether western or non-western morality does not matter here). As I will illustrate later on, the ethical perspective that can propel cultural studies further requires a reflective, participatory, and theoretically informed take.

Given the various meanings that 'culture' has gained within the broad field of cultural studies, it seems almost impossible to provide one simple working definition of the term.[2] In the following, therefore, I will use various concepts of 'body' as analytical lenses to illustrate how culture and normativity can be fruitfully brought together at the intersection of medicine and bioethics. My restriction on this particular intersection has its historical and pragmatic reasons: science and medicine have reconstructed the (late-) modern worldview with regard to

2 This is an almost unavoidable problem that various ethicists and political philosophers struggle with when reflecting on 'culture' from a normative point of view, as Seyla Benhabib 2002 illustrates.

ontology and epistemology and almost replaced the former religious hegemony. They have also fundamentally restructured our everyday life (at least in highly-industrialized societies). Even without any explicit reference to high-tech medicine, human life is from its beginning now structured by modern practices of hygiene, birth control, or prenatal care. The body serves as a locus of all interventions, projects, and expectations. Life expectancy, still an average likelihood, is a mutually shaped result of modern medicine and social conditions, including the composition of expectations and life plans regarding education, family planning, working career, or retirement. That said, the massive impact of other factors such as capitalism, communication technologies, or political orders must be recognized as signifiers of late-modern culture. All of these culminate in the field of medicine. No current debate on medical advancements such as embryonic stem cell research, uterus transplantation, or robotics in health care can restrain the economical, communicative, or legal frameworks in which such debates *and* research practices are embedded.

It seems almost impossible to escape modern medicine's influence from the minute 'we' were created in a pre-birth stage. In a similar vein to the question that Frankenstein's monster rhetorically asks its creator, should we be anxious or thankful about medicine as structural creator of our lives? To escape the emotional stalemate of such a question, I propose to address both the ethical *and* cultural dimensions attached to this matter.

For this purpose, I want to suggest the term of 'elective affinity' (*Wahlverwandtschaft*). It is a productive concept for describing the relationship between normative ethics and socio-cultural studies of medicine that I suggest here. I use the term much as it was used by the German sociologist Max Weber, to describe the fact that two social systems or mentalities are related to or gravitate to each other, even though there is no simple causality or natural logic for such a relationship (see also Swedberg 2005, 83).[3] To construe this relationship as elective affinity is an attractive alternative to the idea that morality and culture are bound by natural kinship. Both the naturalization of cultural values as well as the universal justification attached to social norms neglect the structural differences between the study of norms (the 'ought') with the study of social facts (the 'is'). The elective decision to relate normative judgements to a social practice of morality, embedded in cultural practice, allows for a critical distance to the facts as well as to commonly made claims about how people should behave. The affinity, however, stresses the compatibility of understanding and interpreting social norms and

3 The term *Wahlverwandtschaft* itself stems from early chemistry and was culturally made popular by Johann Wolfgang Goethe's novel *Die Wahlverwandtschaften* 1808.

cultural practice as an expression of moral judgements how things should be. By this, a pragmatic assumption is shining through, meaning that social practice is coined by and therefore expresses moral convictions.

The programmatic shift I suggest by bringing cultural studies and ethical analysis into a productive interplay – intended to allow for an important future turn in cultural studies – intersects on four different dimensions:

First, a mutual critical reflection upon underlying basic assumptions within each area – bioethics as well as cultural studies – is needed. The cultural assumption within applied ethics – here understood as a theoretical reflection on moral practice and everyday norms – often includes limited descriptive conceptions of the self, society, or hegemonic structures. On the other side, the normative premises often buried under the attempt of a 'critical' analysis in cultural studies should be made transparent, visible, and explicit.

Second, on the descriptive-analytical dimension, we need a better, more detailed understanding of the dynamics between biomedicine, lived morality, and socio-cultural factors and how they interact in specific time-space constellations. Here, we are interested on the one side in the processes of negotiations between the somatic, material body defended in biomedicine, and the understanding of the body as locus of cultural inscriptions on the other. Examining this dynamic requires theoretical openness and detailed descriptions of global, local, or glocal developments in the area to enrich our understanding of the complexity.

Third, on a methodological dimension, we should involve lay and patients' moral perspectives beyond the scholars' view. By this, we may appreciate the complexity of the sensing body as promoted by phenomenological or some feminist approaches. Until now, bioethical expertise as well as cultural scholarliness methodologically prioritize the scholar's view on problems, outcomes, and norms. While this, as such is legitimate, it limits our epistemology as well as the range of justifiable claims for generalization. Methodological experimentation and diverse models of inclusiveness need to be addressed as important innovations for the future studies.

Fourth, on a normative dimension, we need to consider integrated approaches to addressing commonalities and parallels in the ethical and cultural space. I suggest the concept of responsibility to increase our analytical sensibility for the political, social body. The 'social body' refers to the power relations defining and ascribing vulnerability, personhood, or injustice related to medical practices. The language and concepts of bioethics need to bring in such concepts for practical and social reasons to overcome the still-unquestioned paradigms of individualism and (neo)liberalism prevailing in bioethics. This opening up of a political-ethical space allows us to rejoin attempts from both sides, from the cultural and the ethical perspectives.

In the following, I will enfold each of these four dimensions. Hereby, I understand each dimension in itself as a field worthy of future research, while the combination of more than one dimension is also welcomed. My programmatic approach takes its self-reflective starting point from bioethics. Regarding the potential of future cultural studies, I do not promote a concrete way to ethicize cultural studies, but I suggest to use the approach I propose here as an analytical lens for disciplinary self-reflection and inspiration. This implies rethinking the underlying assumptions regarding political impacts, the conceptualization of the body, the tendency toward expertocracy and scholarly elitism, as well as issues of responsibility (as scholar, citizen, or society) within cultural studies.

1 Culture and Bioethics: Where to Start

Bioethics is a wide field. On one end of its spectrum, it covers political activities undertaken to implement expert advice (e.g. in form of council or committee). In this context, 'ethics' or 'bioethics' does not mean one clear-cut scholarly way of moral philosophical reasoning; instead, it extends to a broad range of social roles and practical functions. Experts involved are rarely philosophers or ethicists, but can be any kind of academics or legal scholars. Ethicization herein aims at setting up so-called 'soft-law,' often bypassing democratic structures such as parliament or civil society. It presents a governance solution to regulate new social and technological trends (Nowotny, Scott, and Gibbons 2001; Jasanoff 2003). This part of bioethical practice suffers from an underdevelopment of political-ethical theorization as well as from a lack of deliberative and participatory methodologies (Schicktanz, Schweda, and Wynne 2012).

On the other end of its spectrum, bioethics describes a purely scholar activity, based on analytical or sometimes hermeneutic approaches, for developing theories, arguments, or concepts to address ethical problems related to medical practice or life sciences research. As Stephen Toulmin (1982) once put it, this 'applied' context has saved the life of ethics within twentieth-century philosophy. Before then, the area was generally preoccupied with theoretical debates over metaethics and formalistic analytical approaches; and moral philosophy fell victim to this priority. The approach of bioethics as academic endeavor can be characterized by a strong analytical methodology (i.e., considering the moral status of an embryonic stem cell in comparison to a living animal, etc.) or by a narrow focus on very practical questions (i.e., solving clinical ethical conflicts).

A third alternative aims at a cultural and empirically informed bioethics. Apart from all of the challenges and limitations inherent to such an interdisciplinary

enterprise, its real advantages and strengths lie in the integration of intersubjective approaches into a wider context of political-ethical considerations. While conventional bioethics has a strong focus on doctor-patient relationships and on ethical conflicts arising for patients or citizens facing modern science, the cultural and political context (e.g., consumerism and capitalism, Western values and medical ethos, expertocracy and health illiteracy, etc.) in which such a relationship is already embedded is otherwise neglected or disregarded.

Re-contextualizing bioethics as an intellectual activity that acknowledges the political needs and requirements for the public as well as for the academic means to bring back the political-ethical argument. But why was academic bioethics stripped of political-ethical considerations?

This can be explained by at least three factors. First, as a close political-institutional perspective reveals, medicine and life sciences operate mainly outside of parliamentary political structures in many western democracies. While other areas of social life such as trade, work, or education have been highly politicized and heavily regulated since their beginnings, medicine and health care often operate in a rather loose web of political structures. The number of state laws regulating medicine and life science research is rather specific (and are often only a result of public ethical controversies, i.e., as it was the case for embryonic stem cell research, abortion, or organ transplantation). In most western democracies, it is an expertocracy that self-regulates the dos and don'ts by soft-law.[4] Differences between countries exist and it is therefore crucial to study and reflect on the medico-legal culture when examining any particular medical practice and its ethical framework.

A second perspective, in line with a more Foucauldian understanding of 'biopolitics,' acknowledges that there are strong state or institutional interests directly implemented in modern medicine and health care (Lemke 2006). However, they remain implicit and are hidden in the rational language of needs, diagnoses, or treatments. They are already internalized by modern citizens or patients looking after their healthy lifestyles, reproductive behaviors, or end-of-life planning. Such a biopolitical perspective in governance risks eliminating the individual's perspective. The political is all and everywhere, and therefore the productive, analytical tension of the political vs. the non-political is suspended (Bishop and Jotterand 2006). This supra-political perspective might be relevant to understand hegemonic grammar and hegemonic position but it underrates and

4 This argument is supported by the immense impact not only of national academies of science and medicine but also of international organizations such as the World Medical Association or the World Health Organization etc. on the health policy regulation.

oversees practices of resistance or renitence by affected persons (Fraser 1989). Another worry about the domination of biopolitical power as discursive power is that it hides biosociality, unutterable sensation, or embodiment as human factors (Hazan 2015, 27).

A third explanation acknowledges that dominant bioethical approaches such as utilitarian, deontological, or principle-oriented approaches are always embedded in political-ethical assumptions of modern liberalism. However, this political framing became almost invisible because of its presumption of priority given to individualism, to the moral principle of individual choice, and to respect for individual autonomy. This 'naked' version of liberal bioethics suffers from the fact that political assumptions about liberalism entail much more than just this triumvirate. Political-ethical assumptions of liberalism should always include in-depth analyses of the relationship between state, expert, and the citizen; questions of tolerance and its limits; the meaning of collectivity for self-understanding and understanding other's citizens interest; and so on. All these questions surface from time to time in conventional bioethics, but are yet insufficiently addressed. The alternative would be to enlarge the bioethical analysis from the bedside beyond the doors of the hospital: to explore how inter-individual decisions, expectations, and negotiations of lays and professionals are embedded in a broader context of state-market-citizen relationships. Of course, such a zoom is methodologically challenging and limited. However, focusing, for example, on central actors or new political institutions such as patient organizations and patient collectives, allows for such an expanded perspective, which brings together the socio-cultural practices of such collectives (Brown et al. 2004), their political-ethical claims and legitimacy, as well as their impact on bioethical controversies (Beier et al. 2016; Raz, Jordan, and Schicktanz 2014; Schicktanz 2015). Such a normative perspective would complement the cultural study of the collective body – in its explicit as well as more implicit versions – reflecting on the gendered, the disabled, and the colored body, as those bodies are always collectivized.

2 The Body as Local Inscription or as Global Soma: The Dynamics of Medicine, Morality, and Culture

Cultural studies and STS (science and technology studies) have revealed many astonishing facts regarding the dynamics of medicine, cultural practice, and norms. On the one hand, there are areas that can be characterized by strong local

differences or even local resistance against global standards. One example here is the non-acceptance of postmortem organ donation and brain death in Japan and in many other countries of the Asian or Arabic world, while the western world seems to see this as self-evident and taken for granted (Lock 2002; for limits within the western world, see Schicktanz and Wöhlke 2017). On the other hand, there are cases of strong global uniformity and global conformance in medicine, based on assumptions of the body as purely materialistic soma, detached from any interpretation or value (see also Joralemon and Cox 2003).

An example for the global spread of new body technologies is the genetic selection of in-vitro fertilized eggs before they are implanted into a woman's uterus, called as pre-implantation diagnosis (PGD). It is now a commonly accepted practice in most regions of the world across the western/eastern division, if such expensive reproductive medical technologies are affordable. Given the extreme concerns expressed in the early 2000s when PGD was established, its triumphal procession since then is quite impressive and a result of an international active community of scientists and ethicists defending the idea that the fertilized eggs are not yet morally relevant as the 'adult' human body and its attached personhood.

However, there are also more complex examples, of how modern medicine is both globally spreading *and* locally adapted to fit into the respective cultural context. This process can be understood as 'glocalization.' Here, the concept of glocalization is understood to analyze the process of negotiation, refraction, and mimicry between globalization and localization (see Bauman, 1998; Roudometof 2016, 1–42). In contrast to globalization – here understood as the modern version of a market-driven soft-colonialism – glocalization as a conceptual approach sensitizes for a detailed analysis of how the global and the local are negotiated case by case in medical and health care practice. The local-global relationship of various medical practices might differ with regard to their legal-ethical frameworks (for example in the case of organ donation and its different regulations worldwide: Shepherd, O'Carroll, and Ferguson 2014; Lopp 2013; Randhawa and Schicktanz 2013). Economical aspects, regarding when and how much a new medical technology is covered by public health insurance, are also an obvious striving force for global spread. While some public health systems cover all costs for in-vitro fertilization for every woman, others cover a limited number of treatments only for heterosexual couples (Brigham, Cadier, and Chevreul 2013). Even the scientific practices might also differ, for example, in which gynecological examinations are conducted in the US, France, or Germany, as once observed by Lynn Payer (1989). More often it remains globally robust, because common medical diagnoses or treatments are now conducted along international standards to satisfy the quality criteria of the 'gold standard' of the World Medical Association.

To provide a more detailed picture of what I mean by 'glocalization,' I will expand upon the current practice of *surrogacy* as an illustrative example of the way that medical practice, culture, and morality are mutually negotiated. Surrogacy is an artificial reproductive treatment where a so-called surrogate, the gestational mother, is implanted with a genetically often non-related embryo, then, after birth, hands the baby over to the so-called social parents. The surrogate and the intended parents are bound via a contract, and the intended parents normally adopt the child after birth or are legally acknowledged because the embryo is genetically related to them.[5]

The idea to implant a fertilized egg into a womb of a woman not genetically related became technically possible after the introduction of in vitro fertilization, from the end of the 1970s on. Two decades later this practice has spread worldwide (see Mitra, Schicktanz, and Patel 2018, 3–6). While some South Asian regions are often portrayed in the media as hot spots for surrogacy markets, surrogacy is also now practiced in all other continents of the world. However, the concrete practice varies extremely with regard to the selection process of surrogates as well as access for potential parents. The surrogate can include a close relative acting out of 'altruistic' reasons, or an almost unknown person selected from an internet databank. Defense of a commercialized practice of surrogacy sees the surrogate as a 'womb to rent,' and the delivery of a baby as bio-labor, which needs to be reimbursed in 'fair' prices. Alternately, the proponents of 'altruistic' practice assume an emotional bond between surrogate and baby via physical unity, and therefore want to avoid any commercialization or allow bonds between the child and the surrogate. The legal justifications for eligible intended parent(s) differ strongly, too: in India, currently, only heterosexual couples are allowed for medical reasons, while in Israel religious reasons determine who can be a surrogate in relation to the intended parents (e.g., only a Jewish surrogate for Jewish-intended parents). In California, homosexual couples or single (male or female) parents can approach a surrogate as intended parents. This is for conservative reasons in many countries not possible because their sexuality is seen as 'unnatural' or 'immoral.' Moreover, the scientific practice differs among countries according to the selection procedure of fertilized eggs or the absolute number of embryos to be implanted into the surrogate's uterus. In most European countries, one, two, or a maximum of three embryos are permitted for implantation, while in the US or India more are possible, despite the significant increase of medical risks associated with multiple

5 In some cases, sperm or egg or both stem from the social parents, but there are also cases where both, eggs or sperms, are donated by another third party.

pregnancies for the surrogate and for the fetuses. In India, selective abortion is practiced to reduce again the number of fetuses if the intended parents want this (see Mitra and Schicktanz 2016).

These variations are inevitably linked to different ethical and social debates regarding the problem of exploitive market conditions for surrogates in low- and middle-income countries (such as India or Thailand), the right of reproductive freedom for intended parents or surrogates, the question of agency of surrogates under unequal social conditions, and the right of intended parents to select or to not come for the baby. The social concerns might be even more general regarding the impact of such a medical practice on the mundane understanding of kinship and motherhood, gender, or ethnicity – always attached to the body.

While an international overview of the debate offers a broad or even balanced picture of the 'pros' and 'cons,' the question needs to be posed whether national, local discourses are also so broad. They seem rather be dominated by few or selective arguments. Such a cultural taming of the ethical debate must be understood as a result of negotiating between the local and global context in which such debates are evolving and – at least for a particular moment in time – are fixable as culturally 'significant.' From a distant, comparative view there gleam some peculiarities: for instance, Indian sociologists have pointed out, despite critiques of the large economic and caste disparities, that it would be important to see the agency and opportunities for self-determination for surrogates, even in situations of commercial surrogacy (Tanderup et al. 2015). In Germany, the agency of surrogates is rarely considered as leading point but ethicists have emphasized rather the 'best interest of the child' as a criterion of legitimization. Whether this argument results in a permission or moral veto is dependent on how the 'best interest of the child' is then concretely interpreted (see also Wiesemann 2016, 133; Beier and Wiesemann 2013). In the US, various scholars have focused on the social risks of commercialization as it might increase social disparities between races or classes and could lead to the exploitation of poor women or to a racist practice of dismissing non-white women as surrogates or egg donors (Thompson 2005, 66). Again, such points to consider are yet rarely addressed in the German context, although might perhaps in the same way be relevant once the practice is implemented.

While none of these points is made exclusively in any of these three different national discourses, it is striking how some main lines of argument prevail in each context. We need more detailed studies to understand how the bioethical discourse depends on the culture in which it is embedded. Such a descriptive-analytical reflection, however, does not solve the quest for a more rational or universal understanding of moral norms – a project still worth to be defended as an ideal orientation, not as a simple solution.

But indeed, such cultural framing of differences in the discourse[6] has led to some misunderstanding within normative ethics: This misunderstanding assumes that the socio-cultural study of differences in bioethical positions results in "normative relativism" (Schicktanz 2018, 117–119). Normative relativism means that we are not allowed to question each other about the local forms of norm validation and norm hierarchy, because all moral considerations are only locally valid. In a more pervasive form of neoliberal argumentation, such relativism is even used to justify any permissive stance towards new technologies: because nation A, B, or C (e.g. the US, India, or the UK) is doing X (e.g. surrogacy), it would be also acceptable in D (e.g. Germany) to do X. The underlying premise is then that there exists no universally valid argument to forbid it.

However, the here-defended idea of 'post-conventional' bioethics' interest in cultural and social studies of medical practice means nothing more or less than contextualizing the leading moral justification by taking into account the 'real perceived' social conditions in which the respective agents (e.g., patients, doctors, citizens) live, as well as the interpretation given to these living conditions. 'Post' because conventional bioethics neglects any cultural embedding or social factors such as gender, ethnicity, class, etc. to influence bioethical positions. Such a distinction between conventional and post-conventional bioethics might provoke objections, because the generalization does injustice to individual scholars who are already open to interdisciplinary exchanges with sociology and cultural studies. However, it is used here to mark a more general shift in the field without discrediting any of those former approaches. Furthermore, it is important not to conflate post-conventional bioethics with postmodern approaches, as the analytical focus on non-relativistic normative traditions is still viable. In a same vein, it would be productive to reflect on normativity within cultural studies by leaving the conventional paths of anti-normativity or radical constructivism. In such a sense, future cultural studies could try to embed their analysis in the lived experiences of moral reflections – not just considering moral standards and values as taken for granted, but to put more emphasis on the human practices of doubts, concerns, sensing dilemmas, seeking deliberation, and how this is culturally mediated.

6 This might be explained by national law and local regulation, though the law also depends on cultural accounts of what is seen as ethical acceptable or not, see Hansen 2012. According to most philosophers the proper way should be that law follows ethics and not the other way around. However, in political practice, legal regulation is sometimes quicker implemented than a thorough ethical deliberation takes place. Therefore, it is important, from a cultural point of view, to assume a rather complex interplay between law and ethics.

The production of hopes, expectations, or fears is accessible through the study of cultural media by understanding the body as text or as narrative (see Dekkers 1998; Frank 1995; Squier 2004, 20–21). Visual and performative art as well as literature are media that allow access to the emotional dimensions in form of a bodily narrative. Examples for this can be found in the work of the French artist Orlan. She examined in her own body art the relationship of pain, medical surgery, beauty, gender stereotypes, and norms. A very different example is Philip Roth's novel *Everyman* (2006), where he examines aging, dying, and end-of-life-planning, and how they impact the relationship of body, personhood, and narration. Both Orlan and Roth share the attempt to display pain and fear of death by pointing to scars, fragility, dependency, and by narrating a lived body in its particular social, moral, and cultural embedding. And both provide a narration of moral doubts and concerns: where are the advantages of modern medicine, where are limits?

However, it is also necessary to contextualize the moral claims made in a historical course of the discourse. The search for reasons for differences or similarities in arguments and norms – and by this, the transgression of geopolitical boundaries – serves as part of a rationalization of each claim made. This is a main condition for the 'elective affinity' of applied ethics and cultural studies and can be seen as productive future for both disciplines.

The challenge of such an approach is not to lead to 'factual fallacy.' Such a factual fallacy would mean jumping directly from empirical or descriptive findings of how people actually think or how practice currently works to the normative conclusion about how it *should* be. Such a normative positivism must be avoided. Instead, we need a critical assessment of how any moral claim or argument brought forward is culturally embedded in a hegemonic presentation. A transparent strategy for a comprehensible, proper making of a *practical-moral judgment* refers to an uncontroversial understanding of practical-moral judgment as *mixed* judgment. The mix consists of a prescriptive (normative) *and* descriptive (empirical, factual) statement combined, but avoids any crypto normativity.

Let us consider for a moment the above-mentioned example of surrogacy. Consider that somebody states in a public debate that surrogacy should be allowed in Germany, because it allows women a good income and women want this. This claim is a conclusion as practical-moral judgment and built on normative premises (A) *and* on descriptive premises (B). The normative premises can be summarized as the following: A1) Surrogacy is as such morally not wrong and A2) all women should have the right to a good income. Descriptive assumptions that underlie such a conclusion are: B1) Good income is the main interest for women, which presents an empirical question as to whether this is true and women would not value other opportunities to gain more money elsewhere; and B2) It is culturally

uncontroversial what defines 'good income'; or B3) Women have no other chance to get a good income than by surrogacy. What we see by this is that even if we would agree all on A1) or A2), the moral conclusion depends in a paretic version from the descriptive part. This descriptive part requires therefore socially robust knowledge (Nowotny, Scott, and Gibbons 2001, 116–119) about social practice, effects, opinions, etc. Hence, practical-moral judgments do not only depend on common, shared reflections about what is ethically right or wrong, but in a similar way on shared robust interpretations of the world.

3 The Sensing Body as Situated and Affected: Enlarging the Experts' View

Conventional bioethics has a one-sided tendency for the expert discourses. Such a tendency has its historical roots in the analytical tradition of ethics as well as in the close orientation towards the legal discourse. While the analytical approach is not necessarily expertocratic, its formalistic methods and abstract language often hinders non-experts in participation. The legal discourse definitively has an expertocratic manner, given the fact that public education never ever touches upon it and we mainly leave it to specialists, apart from some areas where lay judges are involved. For post-conventional bioethics, the critical assessment of expertocracy is a central element (Schicktanz, Schweda, and Wynne 2012). Cultural insights similarly foster skepticism towards the idea that those not directly affected by or outside of the messiness of everyday life struggles (such as physicians, academic ethicists, or lawyers) can anticipate hypothetically and properly such a complexity in its ambivalence.

Whether this intellectual representation works for the perspectives of persons who are socially marginalized or excluded must be problematized, though, for these persons, as social and political inequality hampers their opportunity to be represented in exclusive circles of academia or other elite groups. Marginalization is here mainly based upon involuntarily, non-mutual membership such as belonging to a particular gender, ethnicity, or nationality. Such a group membership was not voluntarily chosen by these persons, but assigned to them from outside. Marginalization only takes place if a particular group identity is seen as 'negative' (Williams 1998, 15–18). As Melissa Williams has convincingly shown, typical examples of such marginalization have concerned women, people of color, or people with disabilities, depending on particular historical or political-cultural conditions. In relation to such a social exclusion from many intellectual resources or access to socio-political decision making, there is a serious risk that social

stereotypes related to such a negative group assignment hinder those in power to decide to trust testimonials of those from marginalized groups. By this, many public and legal discourses suffer from "epistemic injustice" (Fricker 2007) due to unbalanced power relations in the presentation of knowledge. Countermeasurements include awareness increasing participation or representation by members of one's own group. These very general considerations are particular relevant in the field of medicine (Schicktanz 2015). This is because persons with chronic illness or with a disability have very specific perspectives and insights in the challenges of pain/bodily experiences or social and spatial exclusion when it comes to bioethical issues related to their condition. Being marginalized and affected (meaning that decisions will have a causal effect on them, see Schicktanz, Schweda, and Franzen 2008; Schicktanz 2015 for a detailed definition of 'affected persons') justifies a significant ethical priority of such voices. The embodied or affected experiences as well as the illness identity are legitimate and valuable sources for a privileged understanding of the everyday complexity relevant to the bioethical issue at stake. People in the fourth age, with dementia, or with autism, are too quickly excluded because of the non-translational content of their experiences (Hazan 2015, 47).

Experts, in contrast, cannot phenomenologically rely on such experiences. Of course, they can indirectly reconstruct such experiences by referring to social, cultural, or psychological studies. But finding the right language, the right translation, to transform these special experiences into a social, publicly shared space is not trivial. Hence, the direct involvement of affected persons – in one way or another – is a necessary element for any future bioethics. Because limited resources and basic needs of persons affected might restrict their interest or factual opportunity to take actively part in such discourses and debates, new, joint methods in the cultural and socio-empirical studies can bridge the need for such a reconstruction. Here I see a particular area for future cultural studies to explore various means and methods to bring the affects, interests, vulnerabilities, and needs of those excluded into the broader discourse. In terms of exploration, more anthropological or ethnographic studies of people in the fourth age (Hazan 2015, 46–47, 71) or with dementia are needed to challenge stereotypical and often discriminating views of them as "almost dead" or "cognitive zombies," to enlighten their untypical, but yet human nature. Experimental designs are required to explore the social and ethical issues of biotechnical innovations and their impact on our understanding of humanity. Following the course of cloning novels served Solveig Hansen in her dissertation (2016) as an orientation to examine the historical practice of social othering. By her joint cultural and ethical analysis of how clones have been anticipated and depicted in cultural discourse, she provides a thorough and complex picture of how our moral relationships are

built upon assumptions of sameness in quality (and not in quantity as the clones imply) but also independence and self-reliance as a basis for mutual respect. The limited socio-empirical perspective access to these future scenarios can be productively complemented with such cultural studies of novels and anticipations.

However, this does not mean to incorporate any view of an affected person in an uncritical way. All perspectives shall be reflected with regard to their moral and epistemic claims. Assuming that affected persons are neither able nor willing to transcend their own personal interest into the social sphere is, however, problematic. Emotional as well as biased views are similarly common in experts' debates about patients, persons with disability, or others who are particularly affected. Therefore, any particular position or moral perspective needs to be understood as "situated knowledge" (Haraway 1988) or "situated ethics," but this does not hinder the opportunity to enter a mutual discourse for finding the better argument.

Transferring these thoughts to a future of cultural studies means, for example, to radically revisit the divide of high/classical and pop culture. Especially for any work on medicine and literature, pop culture, such as 'trashy' science fiction or medical thrillers, provide deep insights in common moralities and understandings of modern medicine and biotechnology (see, for instance, Pethes 2005).

4 Body and Responsibility: The Certainty of Moral Tensions as Conjunctions of Deliberation

To illustrate my understanding of post-conventional bioethics as a continuous challenge between practice and theory, between descriptive and normative claims, I want to refer to the performance *Zerreißprobe* (tensile test) of the Austrian artist Günter Brus from 1970. As a performance artist he shocked the public by making his body to the subject of artistic performances. He injured himself by cutting his head and thigh with a razor blade and arranging his vulnerable, naked body half stretched and half hanging within a web of strings crossing a room. Hereby, the vulnerability of the flesh was shown by means of the extreme display of a body disfigured by pain and by interventions from the outside. By being thus displayed, the body itself becomes both the medium of the artistic work, and the scene in which it takes place. It is this mutual meaning and interaction that symbolizes the performance of bioethics by focusing on particular events or single bodies but being aware of the embedding of such entities in a broader context. In a second line of thought, the work of Brus also marks in an abstract sense the particular meaning the 'body' has as intersectional space between bioethics and

cultural studies (see Barkhaus and Fleig 2002, 9–23, 27–36). Although the cultural irritations such artistic presentations of the body produce on their own is not the point here, I assume that almost every viewer of such a performance feels moved, touched, or disturbed. This common sense of vulnerability of the human as body *and* as person is a certainty that provokes the claims of relativism and arbitrariness – Brus's tensile test serves here as a litmus test for the tensions built into modern medicine and biotechnology where they produce, via their innovations, such anticipations of bodily vulnerability and personhood.

So it is precisely the field of body modification and related bioethics where we observe a clash of perspectives in two ways, but which can also serve us as a productive intersection for an elective affinity.

In the first place, there is a serious distinction in the normative ways of ethically judging how we assess the right of self-determination towards our bodies; secondly, there are solid variations in how the body and embodiment are theoretically addressed. I have suggested earlier a methodological approach of making the tension between different meanings of body and autonomy explicit by setting out a dialectical method for heuristic use to be made of the recent dichotomies in bioethics (Schicktanz 2007). By this we cannot easily resolve moral dilemmas, but we can proceed in a dialogical way for addressing theoretically the various descriptive and normative claims. At least, we will overcome simplistic pro- and contra- debates and we are opening up instead of closing down debates for various theoretical relationships between autonomy (and other relevant normative concepts) and body/embodiment. This provides a central interface for the ethical reflection about who can decide what, when, and how about one's own body. What elements of a person can be regarded as available or unavailable at which points in time during the process of this person's life or dying? Whether the 'body boom' in ethics is something avoidable can be questioned (see also Shildrick and Mykitiuk 2005). Even supporters of the liberal conception of self-determination, who primarily recognize the principle of non-maleficence (the general rule not to harm) as morally equivalent, need to clarify the idea of socio-cultural dimensions of embodiment and the framework for the meaning of bodily unavailability within social interaction. Who or what is the other entity that must not suffer damage, and what constitutes damage (Schicktanz 2007)?

Having said this, I need to propel my own focus on the relationship of the normative principle of 'autonomy' on one hand, and the conceptualization of the body/embodiment a bit further. The political-ethical sphere of social interaction requires a constant concern for more than individual autonomy. The most important concepts are then justice and responsibility. Starting from cultural observations and political practice, the bioethical enterprise is not only to set out ideal theories of justice or responsibility but to address witnessed forms of injustice and

irresponsibility. For sure, our sense for such immoralities is neither independent nor free of theoretical presumptions of what justice or responsibilities are. The *ex negativo* start is often more robust regarding our intuitions and knowledge, but it does not free us from a reflective approach to clarify such presumptions.

Considering an opening-up of cultural studies for ethical thinking might be facilitated by art or performative acts that confront us with the limits of textual analysis, rising issues of affects. However, it would be a great self-restriction to limit the ethical perspective to the sphere of aesthetics. Critical sociology and cultural studies studying the presentation or performances of the liminal, excluded, or resistant human existences share a long-standing tradition with concerns about injustice along class, gender, ethnicity, national belonging, or injustice regarding the exclusion of disabled or sick persons. Their arsenal to address injustice is manifold, be it a dense description of exclusion mechanisms or a quantitative summary of the suffering of discriminated parties. They can bring often-unheard voices into the discourse and highlight the agency of parties often neglected or denied: women, children, the ill, or others often overseen (de Beauvoir 2000). This sociological practice, according to Wayne Brekhus (1998), devotes greater epistemological attention to "politically salient" and "ontologically uncommon" features of social life. Addressing women, the elderly, homosexuals, etc., means "marking" those excluded entities, but this practice unreflects or even repeats the hegemonic grammar and leaves the "unmarked" (whites, heterosexuals, men, etc.) unrevoked. Brekhus' critique of the epistemic practice of identity labels and problematic singling-out is important and highly relevant to overcoming simplistic, unreflective assumptions of the good and bad guys. We need to acknowledge that this epistemic practice within sociology is already embedded in a normative theory of justice and fair treatment of which ethics can help to unmask them in future co-operations.

I have suggested somewhere else (Schicktanz and Schweda 2012) that the concept of 'responsibility' is particularly helpful in linking everyday languages of morality and ethical-normative reflection. Providing a theoretical formula for what the concept of responsibility entails offers a way to explicate moral claims of self-responsibility, social or professional responsibility, or family responsibility ubiquitous in medical practice, health policy, or health communication. While of course only working in limitations, the concept of responsibility is not just a moral idea among others, but as a meta-ethical concept, it provides a meaning of how ethical properties are formulated, logically expressed, and epistemically assessed. Therefore, using terms of responsibility means that we are explaining normative claims embedded in social presumptions about relationships. This helps to translate ordinary folk language into a more abstract form to proof for consistency or to detect contradictions.

This need for translation – from the everyday moral grammar to the theoretical -analytical level and back – is an endeavor that goes beyond the conventional understanding of 'education.' It requires theoretical sensitivity for what constitutes a responsible social relationship: It is always embedded in space and time and the relationship is enriched by cultural assumptions (i.e., in the case of the earlier mentioned surrogacy: what constitutes good parenthood; who counts as morally relevant actor: only the intended parents, or also the mother, the doctors, the government, etc.)?

Responsibility, however, in current sociology – especially medical sociology – has been narrowed down to a critical notion of moral imperatives, synonymous with the social practice of blaming and shaming (Rose 2006, 4; Arribas-Ayllon, Featherstone, and Atkinson 2011), and applied to criticize biopolitical strategies. As such, the sociological notion of 'responsibilization' emphasizes a very special application of the term 'responsibility' focusing on the individual or the family as both the moral agent and the moral object in biopolitics. As an alternative, the productive junction with ethical theory alludes that this application has a strong tendency to reduce the understanding and practical usage of responsibility and that there are better, more refined ways to address responsibility in its many dimensions by using a detailed, transparent description of normative complexity (Schicktanz 2016).

5 Summing Up: An Elective Affinity between Bioethics and Socio-Cultural Studies of Medicine and Life Sciences

I have suggested in the beginning the concept of 'elective affinity' to bring forward a new relationship between bioethics and cultural studies – and would mean somehow a double turn-over: a cultural turn for bioethics and an ethical turn for cultural studies. Whether cultural studies have already adopted such an ethical gaze, I am not sure. However, a current trend to differentiate 'critical' cultural studies can be read as tendency to explicitly address issues of marginalization, discrimination, and exclusion. Being informed by various strands of critical theory might, however, not be the only future direction for cultural studies. Other approaches stemming from applied bioethics to address various ethical and social concerns can be innovative and helpful, as suggested here.

Normative studies and moral languages provide access to moral practice and help to signify the consistency as well as inconsistency in moral practice and

ethical thinking. There exists no absolute demarcation between moral practice in everyday life and theoretical ethical reflection, rather it is a continuum with smooth transitions. The theoretical concept of "reflective equilibrium" serves as dialectical model of normative judgment between theory and practice to describe and reflect on this continuum as philosophical method (Daniels 1996).

It is, however, seen as a legacy of Max Weber's idea of value-free sociology and economical sciences that until today lead to an unrealistic or even wrong ideal of value-free sociology or cultural studies. As the dispute over 'value-free' (*Wertfreiheit*) vs. 'value judgment-free' (*Werturteilsfreiheit*) revealed already almost a half century ago (Albert and Topitsch 1971), it is not only a strange mythos of modern social and cultural studies to be value-free, as the study of values as well as the explication of values is part of any scientific activity – in social and cultural sciences, as well as life science areas such as medicine or agriculture aiming for 'saving life,' 'curing disease,' or providing 'better living conditions' or 'sustainability.' However, such scholars should be explicit and transparent when making value judgments instead of allowing crypto-normativity in scientific terms or scholarly language. Terms and concepts such as 'critique,' 'power,' 'social inequality,' 'vulnerability,' and 'colonialism' always reflect a pejorative, moral meaning that we cannot escape as either speaker nor listener (Fraser 1989, 17–20). However, not any value judgment can claim to count as well-considered judgment. Applied ethics and moral philosophy provide the methodological arsenal to win this battle over crypto-normativity and hidden values in scholarly work.

A flourishing, productive elective affinity between bioethics and cultural studies requires a crucial clarification about all own normative premises on the why, the how, and the what of ongoing research. The 'why'-question focuses on the motivations and programmatic reasons behind a study and for singling out a problem to being relevant for in-depth examination. The 'how'-question follows the lines of a chosen methodology and asks how far normative premises are already embedded in the research program (Merton and Storer 1973, 229–250). For example, does the selection of qualitative vs. quantitative methods only refer to epistemic assumptions of generalization or depth, or might it also include who should be in the focus of examination (the lay public, the experts, the media, etc.). The 'what'-question critically reflects which underlying assumptions of injustice, responsibility, or vulnerability are already attached to the selection of a particular topic (the topic of terror in Europe, the topic of dementia in India, the topic of education in Africa, etc.). How does the spot on this topic risk shading other topics, and is the priority well-justified?

From a bioethicist's point of view, there are many good arguments as to why and how socio-cultural studies are important or even indispensable for a well-defined and well-argued problem definition (what is the moral problem we

want to solve) as well as the practical recommendations often following bioethical inquiries 'how to solve' the conflict in the future. The concrete function of socio-cultural empirical insights for norm justification is instead very controversial and perhaps for the purpose of the here-proposed collaboration modus not needed.

Cultural studies provide not only, but still importantly, a challenge to monological or one-sided perspectives on bioethics. From a theoretical point of view, the solution to the problem of legitimacy lies not in simple forms of public participation in research and policy making, but in a conceptual analysis of the kind of perspectives needed. I am here assuming that there is no single, ultimate perspective. Only a combination and pluralization of different perspectives can offer us an approximation of the 'whole picture.' This requires a systematic adoption of other perspectives (Schicktanz 2015, 251–252). With this increased complexity, we enhance our understanding of the dependence of morality on affects and social dimensions of power. Thinking with stories, narratives, or images as cultural studies provides the arsenal and methodology that help us to test for consistency, for the wideness of the chosen perspective, or for the peculiarity of it. However, there are new risks such as more hidden morality, exclusivity, and ambiguity awaiting such concerted efforts. It would be worth going still further in this direction in the future.

Acknowledgement: I am particularly grateful to Haim Hazan, Tel Aviv, and Solveig L. Hansen, Göttingen, for critical-constructive comments to an earlier version, and would also like to thank Doris Bachmann-Medick and Jens Kugele, Giessen, for their editorial feedback.

References

Albert, Hans, and Ernst Topitsch, eds. *Werturteilsstreit*. Darmstadt: Wissenschaftliche Buchgesellschaft, 1971.

Arribas-Ayllon, Michael, Katie Featherstone, and Paul Atkinson. "The Practical Ethics of Genetic Responsibility: Non-Disclosure and the Autonomy of Affect." *Social Theory & Health* 9.1 (2011): 3–23.

Barkhaus, Annette, and Anne Fleig, eds. *Grenzverläufe. Der Körper als Schnittstelle.* Munich: Fink, 2002.

Bauman, Zygmunt. *Postmodern Ethics*. Hoboken, NJ: Wiley, 1993.

Bauman, Zygmunt. "On Glocalization: or Globalization for some, Localization for some Others." *Thesis Eleven* 54.1 (1998): 37–49.

Beier, Katharina, and Claudia Wiesemann. "Reproduktive Autonomie in der liberalen Demokratie – Eine ethische Analyse." *Patientenautonomie. Theoretische Grundlagen –*

Praktische Anwendungen. Eds. Claudia Wiesemann and Alfred Simon. Münster: Mentis, 2013. 199–216.
Beier, Katharina, Isabella Jordan, Claudia Wiesemann, and Silke Schicktanz. "Understanding Collective Agency in Bioethics." *Medicine, Healthcare and Philosophy* 19.3 (2016): 411–422.
Benhabib, Seyla. *The Claims of Culture: Equality and Diversity in the Global Era*. Princeton: Princeton University Press, 2002.
Bishop, Jeffrey P., and Fabrice F. Jotterand. "Bioethics as Biopolitics." *Journal of Medicine and Philosophy* 31.3 (2006): 205–212.
Brekhus, Wayne. "A Sociology of the Unmarked: Redirecting Our Focus." *Sociological Theory* 16.1 (1998): 34–51.
Brigham, Karen, Benjamin Cadier, and Karine Chevreul. "The Diversity of Regulation and Public Financing of IVF in Europe and its Impact on Utilization." *Human Reproduction* 28.3 (2013): 666–675.
Brown, Phil, Stephen Zavestoski, Sabrina McCormick, Brian Mayer, Rachel Morello-Frosch, and Rebecca Gasior Altman. "Embodied Health Movements: New Approaches to Social Movements in Health." *Sociology of Health & Illness* 26.1 (2004): 50–80.
Brus, Günter. *Performance Art: Zerreißprobe [Tensile Test]*. https://www.youtube.com/watch?v=Gmtm_A5fvbc. Munich, 1970 [accessed: 1 July 2019].
Clifford, James, and George E. Marcus, eds. *Writing Culture: The Poetics and Politics of Ethnography.* Berkeley/Los Angeles/London: University of California Press, 1986.
Daniels, Norman. *Justice and Justification: Reflective Equilibrium in Theory and Practice.* Cambridge: Cambridge University Press, 1996.
De Beauvoir, Simone. *Das andere Geschlecht. Sitte und Sexus der Frau*. Reinbek: Rowohlt, 2000.
Dekkers, Wim. "The Body as a Text: The Interpretive Tradition." *Advances in Bioethics*. Eds. Rem B. Edwards, Edward E. Bittar, and Martyn Evans. Stamford, CT: JAI Press Inc., 1998. 209–228.
Douglas, Mary. *Natural Symbols: Explorations in Cosmology.* London: Barrie & Rockliff, 1970.
Douglas, Mary. *Risk and Blame: Essays in Cultural Theory.* London/New York: Routledge, 1992.
Frank, Arthur W. "Illness as a Call for Stories." *The Wounded Storyteller: Body, Illness, and Ethics.* Ed. Arthur W. Frank. Chicago/London: University of Chicago Press, 1995. 53–74.
Fraser, Nancy. *Unruly Practices: Power, Discourse and Gender in Contemporary Social Theory.* Minneapolis: University of Minnesota Press, 1989.
Fricker, Miranda. *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford/New York: Oxford University Press, 2007.
Geertz, Clifford. *The Interpretation of Cultures: Selected Essays*. London: Hutchinson & Co, 1973.
Goethe, Johann Wolfgang. *Die Wahlverwandtschaften. Ein Roman.* Berlin: Insel (2012 [1808]).
Hansen, Solveig L. "Utopische Dynamiken von Macht, Reproduktion und Gesetz? Ein interdisziplinärer Beitrag zur Law-and Literature-Debatte." *Gehört mein Körper noch mir?* Ed. Susanne Beck. Würzburg: Nomos, 2012. 193–221.
Hansen, Solveig L. *Alterität als kulturelle Herausforderung des Klonens. Eine Rekonstruktion bioethischer und literarischer Verhandlungen.* Münster: Mentis, 2016.
Haraway, Donna. "Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective." *Feminist Studies* 14.3 (1988): 575–599.
Hazan, Haim. *Against Hybridity: Social Impasses in a Globalizing World.* Cambridge/Malden, MA: Polity, 2015.

Jasanoff, Sheila. "Technologies of Humility: Citizen Participation in Governing Science." *Minerva* 41.3 (2003): 223–244.
Joralemon, Donald, and Phil Cox. "Body Values: The Case Against Compensating for Transplant Organs." *The Hasting Center Report* 33.1 (2003): 27–33.
Lemke, Thomas. "Genetic Responsibility and Neo-Liberal Governmentality: Medical Diagnosis as Moral Technology." *Michel Foucault and Power Today: International Multidisciplinary Studies in the History of the Present.* Eds. Alain Beaulieu and David Gabbard. Lanham, MD: Lexington Books, 2006. 83–91.
Lévi-Strauss, Claude. *Tristes Tropiques.* London: Hutchinson & Co, 1961 [1955].
Lévi-Strauss, Claude. *Structural Anthropology*. New York: Basic Books, 1963 [*Anthropologie Structural*, 1958].
Lock, Margaret. *Twice Dead: Organ Transplants and the Reinvention of Death.* Berkeley/Los Angeles/London: University of California Press, 2002.
Lopp, Leonie. *Regulations Regarding Living Organ Donation in Europe: Possibilities of Harmonisation.* Berlin/Heidelberg: Springer, 2013.
Merton, Robert K., and Norman W. Storer. *The Sociology of Science: Theoretical and Empirical Investigations.* Chicago: University of Chicago Press, 1973.
Mitra, Sayani, and Silke Schicktanz. "Failed Surrogate Conceptions: Social and Ethical Aspects of Preconception Disruptions During Commercial Surrogacy in India." *Philosophy, Ethics, and Humanities in Medicine* 11.9 (2016) https://peh-med.biomedcentral.com/articles/10.1186/s13010-016-0040-6 [accessed: 1 July 2019].
Mitra, Sayani, Silke Schicktanz, and Tulsi Patel. *Cross-Cultural Comparisons on Surrogacy and Egg Donation: Interdisciplinary Perspectives from India, Germany and Israel*. London: Palgrave Macmillan, 2018.
Nowotny, Helga, Peter Scott, and Michael Gibbons. *Re-Thinking Science: Knowledge and the Public in an Age of Uncertainty.* London: Polity, 2001.
Payer, Lynn. *Culture and Medicine*. London: Penguin, 1989.
Pethes, Nicolas. "Terminal Men: Biotechnological Experimentation and the Reshaping of 'the Human' in Medical Thrillers." *New Literary History* 36.2 (2005): 161–185.
Randhawa, Gurch, and Silke Schicktanz, eds. *Public Engagement in Organ Donation and Transplantation*. Lengerich: Pabst Science Publishers, 2013.
Raz, Avid, Isabella Jordan, and Silke Schicktanz. "Exploring the Positions of German and Israeli Patient Organizations in the Bioethical Context of End-of-Life Policies." *Health Care Analysis* 22.2 (2014): 143–159.
Rise of the Planet of the Apes. Dir. Rupert Wyatt. Narr. Rick Jaffa and Amanda Silver. 20th Century Fox, 2011.
Rose, Nikolas. *The Politics of Life Itself: Biomedicine, Power, and Subjectivity in the Twenty-First Century*. Princeton: University Press. 2006.
Roth, Philip. *Everyman.* New York: Houghton Mifflin, 2006.
Roudometof, Victor. *Glocalization: A Critical Introduction*. London/New York: Routledge, 2016.
Schicktanz, Silke. "Why the Way We Consider the Body Matters: Reflections on Four Bioethical Perspectives on the Human Body." *Philosophy, Ethics, and Humanities in Medicine* 2.30 (2007): 1–12.
Schicktanz, Silke. "The Ethical Legitimacy of Patients Organizations' Involvement in Politics and Knowledge Production: Epistemic Justice as Conceptual Basis." *The Public Shaping of Medical Research: Patient Associations, Health Movements and Biomedicine*. Eds. Peter Wehling, Willy Viehöver, and Sophia Koenen. London/New York: Routledge, 2015. 246–265.

Schicktanz, Silke. "Genetic Risk and Responsibility: Reflections on a Complex Relationship." *Journal of Risk Research* (2016): 1–23.
Schicktanz, Silke. "Beyond Relativism: Comparing Practices and Norms of Surrogacy in India, Israel and Germany." *Cross-Cultural Comparisons on Surrogacy and Egg Donation: Interdisciplinary Perspectives from India, Germany and Israel*. Eds. Sayani Mitra, Silke Schicktanz, and Tulsi Patel. London: Palgrave Macmillan, 2018.103–125.
Schicktanz, Silke, Mark Schweda, and Martina Franzen. "'In a Completely Different Light?' The Role of 'Being Affected' for the Epistemic Perspectives and Moral Attitudes of Patients, Relatives and Lay People." *Medicine, Health Care and Philosophy* 11.1 (2008): 57–72.
Schicktanz, Silke, and Mark Schweda. "The Diversity of Responsibility: The Value of Explication and Pluralization." *Medicine Studies* 3.3 (2012): 131–145.
Schicktanz, Silke, Mark Schweda, and Brian Wynne. "The Ethics of 'Public Understanding of Ethics': Why and How Bioethics Expertise Should Include Public and Patients' Voices." *Medicine, Health Care and Philosophy* 15.2 (2012): 129–139.
Schicktanz, Silke, and Sabine Wöhlke. "The Utterable and Unutterable Anthropological Meaning of the Body in the Context of Organ Transplantation." *Dilemata: International Journal of Applied Ethics* 9.23 (2017): 107–127.
Searle, John. *The Construction of Social Reality*. New York: Free Press, 1995.
Shelley, Mary. *Frankenstein.* New York: Dover Publications, 1994 [1818].
Shepherd, Lee, Ronan E O'Carroll, and Eamonn Ferguson. "An International Comparison of Deceased and Living Organ Donation/Transplant Rates in Opt-in and Opt-out Systems: A Panel Study." *BMC Medicine* 12 (2014): 131. doi:10.1186/s12916-014-0131-4 [accessed: 30 June 2019].
Shildrick, Margrit, and Roxanne Mykitiuk, eds. *Ethics of the Body*: *Postconventional Challenges*. Cambridge, MA/London: MIT Press, 2005.
Squier, Susan M. *Liminal Lives: Imagining the Human at the Frontiers of Biomedicine*. Durham, NC: Duke University Press, 2004.
Swedberg, Richard. *Max Weber Dictionary: Key Words and Central Concepts*. Stanford, CA: Stanford University Press, 2005.
Tanderup, Malene, Sunita Reddy, Tulsi Patel, and Brigitte Bruun Nielsen. "Reproductive Ethics in Commercial Surrogacy: Decision-Making in IVF Clinics in New Delhi, India." *Journal of Bioethical Inquiry* 12.3 (2015): 491–501.
Thompson, Charis. *Making Parents: The Ontological Choreography of Reproductive Technologies.* Cambridge, MA: MIT Press, 2005.
Toulmin, Stephen. "How Medicine Saved the Life of Ethics." *Perspectives in Biology and Medicine* 25.4 (1982): 736–750.
Wiesemann, Claudia. *Moral Equality, Bioethics, and the Child.* Cham: Springer, 2016.
Williams, Melissa. *Voice, Trust, and Memory: Marginalized Groups and the Failings of Liberal Representation.* Princeton: Princeton University Press, 1998.

Normativity, system-integration, natural detachment and the hybrid hominin

Lenny Moss

Phenomenology and the Cognitive Sciences

ISSN 1568-7759

Phenom Cogn Sci
DOI 10.1007/s11097-020-09682-2

Springer

Phenomenology and the Cognitive Sciences
https://doi.org/10.1007/s11097-020-09682-2

Normativity, system-integration, natural detachment and the hybrid hominin

Lenny Moss[1]

Published online: 19 June 2020

Abstract

From a subjective point of view, we take the existence of integrated entities, i.e., ourselves as the most unproblematic given, and blithely project such integrity onto untold many "entities" far and wide. However, from a naturalistic perspective, accounting for anything more integral than the attachments and attractions that are explicable in terms of the four fundamental forces of physics has been anything but straightforward. If we take it that the universe begins as an integral unity (the singularity referred to as "the cosmic egg") and explodes into progressive stages of internal detachment, then we can also fathom the idea that eddies of relative detachment becoming increasingly integral. Helmuth Plessner made a powerful case for the onset of "positionality" constituting one of the major transitions in nature. Surely, the emergence of "entities" (i.e., life-forms) that position themselves in relation to their surround marks a decisive transition in relative levels of detachment and some would say "autonomy." It would follow, with no less force, that where and when entities can be seen to be normatively integrated, and indeed to be the agents of their own normativity, that another threshold of detachment has been crossed. The paper introduces and explores the idea that normativity, embedded in a wide-ranging theory of natural detachment, can be considered an emergent force of nature that is requisite to accounting for levels of integration beyond that which is explicable in terms of the four fundamental forces of physics. Following this line of enquiry, we argue that the first expression of a fully, normatively-integrated life-form is neither a spoken language user nor for that matter an individual but rather the neoteny-based, Homo erectus Group. In so doing we claim to have made an inroad into embedding the force of normativity into a wide-ranging naturalist framework, to have provided philosophical anthropology with a new (post-individualist) point of departure, and at least playfully, to have given some naturalistic grist to Hegel's proclamation that spirit (Geist) is the truth of nature.

Keywords Normativity · Detachment · Hybrid hominin · System integration · Homo erectus · Plural-self awareness · Cerebral asymmetry · Mimesis

✉ Lenny Moss
dcimorse@icloud.com

[1] Department of Sociology, Philosophy and Anthropology, University of Exeter, Exeter EX44RJ, UK

1 Natural integration and natural normativity

The problem of natural integration, or perhaps unification – the constitution of a unity – is a truly hard problem that has seldom if ever been addressed as such. Indeed, it is perhaps *the* hard problem of the philosophy of life and mind. Granted, how a natural entity can have interiority, i.e., subjectivity, is a hard problem, but if the question of integration and unification is not identical to the "Hard Problem of Consciousness" (Chalmers 1995) it is also inseparable from it and surely a presupposition of the very possibility of interiority. Nor would even an understanding of how subjectivity could be resident to a single cell tell us how consciousness could become an integrated unity across many cells. If we assume consciousness is a physically based phenomenon, and that it draws upon the activity of various parts of the brain, let alone constituent cells, then we must face our deficits in understanding how the experience of a unified consciousness is realized at the level of an integration of some cells but not others albeit in the absence of evident, non-arbitrary, physical boundaries.

Both more fundamental than the Hard Problem of Consciousness, and more expansive in scope, the problem of integration/unification is also central to the problem of the origin(s) of life. If merely conventionally mechanistic associations of parts outside of parts (*partes extra partes*)[1] were sufficient for creating a living unit, it would have been achieved long since. The fact that a natural bacterial chromosome could be successfully replaced by an artifactual chromosome only further testifies to the irreducible importance of a naturally integrated unit, the bacterial cell, as the prior condition of possibility for any molecular configuration, natural or artificial, to *be* a chromosome.[2] Perennial claims by biologists to have solved the problem of the origins of life have either been projections based upon the largely uncritical metaphysics of informational idealism (masquerading as consensual science), or based on the largesse of liberal promissory notes fueled by commercial marketing interests.

Lest there be any doubt, I do not aspire to overcome the problem of integration *tout court* in this paper. Beyond thematizing it as a problem (a worthy endeavor in itself, I claim) I will want to explore the role of "normativity" in hominin/human integration and toy with the idea of normativity being *a force of nature*. Clearly doing so, if feasible at all, would require a radical expansion, along with a capacity for scaling, of our current concept of normativity. I hope to at least provide some adumbration of what that might look like.

For Leibniz, Kant and Hegel (among others) normativity plays a central role in the understanding of life and mind, albeit in different ways. Hegel, for our present *naturalistically* oriented purposes, is of most interest because he comes the closest to offering an account of the transition between a pre (or perhaps proto?) normative

[1] I first heard the wonderfully mellifluous phrase "partes extra partes" when uttered by Charles Taylor in lectures, possibly his lectures on Hegel at Berkeley in 1981. Taylor himself took the phrase from Merleau-Ponty.

[2] When Craig Venter announced that he had created artificial life in 2010 it resulted in splash headlines in media outlets as respectable as The Guardian and convinced lay readers that Venter had succeeded in creating life de novo. https://www.theguardian.com/science/2010/may/20/craig-venter-synthetic-life-form. The project cost 40 million dollars and was addressed by the responsible investigators themselves as representing an advance in the technology of synthetic biology, not in terms of relevance to questions of the origins of life (Gibson et al. 2010).

account of Nature and a fully normative account of Spirit. As Brandom tells us: "*How one understands the relation between these, both conceptually and historically, is evidently of the first importance in understanding what Hegel has to teach us about the normative realm he calls 'Geist'*" (Brandom 2019).[3]

Hegel, as most clearly and emphatically presented by Brandom, provides a fundamental account of how the very glue that enables the human realm of culture to exist, that integrates and unifies it, is the glue of normativity. Normativity on this level is about a force which constitutes and unifies human life, albeit not a force that can be reckoned in terms of the vocabulary of classical mechanics. Human life entails the existence of self-conscious subjects who are constitutively enmeshed in the force field of reciprocal accountabilities and entitlements. To be a self-conscious subject, is to have a self-identity, to be not just an entity in-itself but an entity for-itself. An entity for-itself doesn't exist in a vacuum. To be an entity for-itself is an ongoing social achievement dynamically realized within the normative fabric of human life. As George Herbert Mead argued, a 'self' is a reflective objectification that we discover from a certain social vantage point. There is the self we are as son or daughter, as husband or wife, as father or mother, as teacher or carpenter, and so on. We are neither the passive residue of forces out of our control nor the sovereign dictators of the self we are, we are active in taking stands in each social location about who we are, but our ability to find ourselves to be that self is mediated by the constructive recognition of others. All the possibilities of being a certain self are normatively defined (they are located with the framework of entitlements and accountabilities). To be a being-for-itself is thus an on-going accomplishment achieved necessarily within the fabric of social-normative space.[4]

Within the fabric of normatively structured reciprocal accountabilities and entitlements we take a stand on whom we claim we are entitled to be recognized as. And we do so in the context of reciprocally recognizing those whose recognition we require. To speak of entitlements, obligations and accountabilities is to speak of forces (or *power* in the Foucaultian sense). Hegel famously refers to the *I that is a We and the We that is an I*.[5] It can only be a We that enables me to be an I, and thus reciprocally a We can only be a We if it is situated by an I. The obligate relation of I and We describes a special kind of palpable integration and unity. Hegel has done original and foundational work with respect to characterizing the kind of normatively powered unity and integration that constitute the phenomenon of Geist (only in the context of which selves can be selves). But that is not to say that Hegel offers a full-fletched, naturalistic answer to the problem of integration as we've defined it.

For both Leibniz and Kant, normatively resonant entities must necessarily be presupposed and cannot be theoretically derived from pre-normative precursors. As

[3] Brandom's much anticipated tome on Hegel had not yet be released as of this writing. The quotations are derived from his 2014 draft of the text made available on his home page https://www.pitt.edu/~brandom/spirit_of_trust.html. The quotation above is from A Spirit of Trust: A Semantic Reading of Hegel's Phenomenology, Part Three Self-Consciousness and Recognition, Chapter Seven: The Structure of Desire and Recognition: Self-Consciousness and Self-Constitution.

[4] For an excellent discussion of Mead's concepts of I and self and of the intersubjective origins of self-hood see Habermas 1992.

[5] Hegel, GWF, Phenomenology of Spirit, (trans. A.V. Miller) Oxford: Oxford University Press, section 177, 110.

fundamental substance Leibniz's monads are as basic as it gets and are always already normatively-infused fully integrated units. Only God could provide the recipe for how to make one. Likewise, the Critical Philosophy holds in principle that neither the origins of the ostensibly purposive functional and organizational unity of any living being, nor certainly that of the normativity of pure practical reason that infuses "Man's moral vocation", could ever be accounted for by (human) theoretical means.

Hegel, by contrast, appears to provide a developmental account in which animal desire becomes an entryway into Geist when desire becomes a desire for recognition. But even the assumption of animal desire is already assuming a unity of purpose that takes the fundaments of what I am calling the really hard problem as an assumed given. Hegel has at least offered the prototype of a dialectical account of movement from a lower to a higher level of unity. Inasmuch as Hegel is, as his famous title suggests, performing a phenomenology of mind, some manner of mindedness is always already assumed. Hegel is thus not accountable to the kind of really hard problem that a contemporary naturalism would pose as it would not be immanent to his phenomenological standpoint. The Absolute, and the Concept, which are ontological primitives for Hegel, supersede the problem of integration and unification in advance, albeit by phenomenological fiat.

Neither a phenomenological nor a transcendental philosophy alone can offer a fully naturalistic account of the basis of integration and unity, given the constitutively methodological assumptions and constraints of each approach. And this is no less the case for Heidegger's existential phenomenology that takes its point of departure from *Dasein's* practical involvement in her normatively integrated world. But nor is a reductive empiricism the answer to going beyond the limits of a *partes extra partes* vision of nature. Expressed in very simple terms, neither a science that limits itself entirely to a bottom-up approach, nor a science that limits itself to only a top-down approach, will be able to solve the problem of integration and unification at all levels of nature. What we require is an approach that can bring both of these together in some way. In recent work Arran Gare has made a compelling case for how this involves a renewed acceptance and embrace of speculative thinking (Gare 2018). For physicist Lee Smolin and philosopher Roberto Mangabeira Unger (2015), whose collaborative work over eight years has resulted in a fascinating book entitled *The Singular Universe and the Reality of Time*, bringing together the big picture of the universe as a whole with the micro-level advances of physics, entails no less than a return to the practice of Natural Philosophy proper. In referring to "normativity as a force of nature", I have thus far only used as a teaser, or perhaps appetizer, the idea that normativity might be able to span the chasm between the bottom up and the top down, and thus serve as a conceptual lynchpin of a renewed speculative natural philosophy. There will be more to come. But it should also be acknowledged that attempting to bring together a top-down with a bottom-up approach, even to also address questions of integration and unification is not entirely new.

Just shy of 100 years ago, Helmuth Plessner, who had trained in both biology and philosophy, put forward the theme of "positionality" as the key concept in an attempt to weave together empirical and phenomenological, bottom-up and top-down, big picture and micro-detail, perspectives into a coherent natural philosophy and philosophical anthropology. Plessner included fundaments about biological structure, organization,

physiology, ecology, and ethology in his attempt at elaborating a non-dualist, and systematically coherent framework for all forms of life (Plessner 2019).

Using positionalilty as a point of departure for elaborating a logic of life, including human life, is in fact tantamount to simultaneously treating the question of integration, or unification, as a point of departure. As Plessner makes clear, it is only with the living state that boundaries become something other than arbitrary and contingent. To the extent that life is defined by anything, it is defined by the dynamic enactment of a distinction between inner and outer, Plessner's so-called "double aspectivity." "Positionality" then is the categorical universal and sine qua non of the living state and conceptually defines the boundary between physical phenomena that, for all intents and purposes, are *partes extra partes,* from those for which some non-trivial level of system unity and integrity has been achieved. Plessner does not specifically take up the language of *normativity* and yet one can find it to be implicit in his account. The realization of a life, even at the most rudimentary, let's say the simple, single-celled level, already entails a form of active mediation between reaching outward and enforcing a boundary, that suggests the regulatory enactment of an implicit norm. For Plessner, there is a logic of dialectical building of positional levels upon levels of reflection that culminate in the human level of "excentric positionality" whereby the "shared" (and invariably normative) perspective of the *universal Other* is always reflexively embodied and reflectively available. Which is to say that Plessner has long since offered a body-mind neutral account in which human-level normativity is located on a natural continuum in which questions of dynamic system integration are at least implicitly fundamental. Where Plessner fell short, I suggest, is in (only) deriving a largely monological account of the emergence of human-level normativity.[6] The following may be viewed as in part an attempt to offer the complementary perspective, albeit with the full reconciliatory and synthetic engagement to appear in subsequent work.

2 Natural detachment and the hybrid hominin

2.1 The idea of detachment

The idea of "detachment" may seem on first pass as a very odd way to begin a discussion about the basis of integration. However, if we begin with some notion of the universe as a whole in a state of a kind of primordial integration then perhaps we can fathom how detachments from a simple and primordial integration would be a pre-condition for new and perhaps more interesting and intricate forms of integration. When we take the human self (however "natural" doing so may feel) as an unproblematic given we quickly obscure the issue. No matter how you cut it, from any naturalistic point of view, the self is highly derived state of affairs. We can't afford to forget this lest we risk inadvertently tumbling back into metaphysical dualism. There will be a fairly compelling case to be made for the relevance of the idea of detachment

[6] CF. my preliminary account of this criticism in my Notre Dame Philosophical Reviews discussion of the recent English translation (Plessner 2019) https://ndpr.nd.edu/news/levels-of-organic-life-and-the-human-an-introduction-to-philosophical-anthropology/.

to the transition (or transitions) constitutive of hominin evolution and indeed this idea goes back to the late eighteenth century. The more speculative mood deigns to propose that detachment can be scaled and tracked *all the way down*. For present purposes the objective is not to aspire for the idea of detachment to assume the status of a new master theory but rather to lobby for its value in a perspectival sense and in so doing simultaneously affirm the wisdom of some measure of perspectivism in our thinking.

The intuition I wish to arouse is that an entity that claims some measure of detachment is to that extent standing on its own feet, and the greater the level of detachment the more self-standing it is. Detachment then would also suggest internal unity and thus integration, and so must be a kindred concept, and yet detachment is not identical to integration. It can't be identical with integration because, as we will see, there is also a special kind of detachment which is "downward" and disintegrative. More on this later.

Detachment, we will have to assume, begins at the beginning, i.e., with the Big Bang. Prior to the Big Bang, all existence (whatever that means) was confined to an infinitely dense, infinitesimally small singularity in which there was no space or time and all four basic forces of nature (as we now understand them) were united into one. The universe was born, space-time emerges, in an explosion of detachment. Without getting too bogged down in the technical details of high-energy physics and cosmology, the take-home lesson is that a logic of detachments-built-upon-detachments is set into motion from... the beginning. It has been theorized that cosmic detachment began with the detachment of gravity from the unity of fundamental forces, resulting in the formation of elementary particles and anti-particles followed by an *inflation* into space-time triggered by the further detachment of the strong nuclear force. As yet inexplicable asymmetries in the appearance of baryons (matter) versus antibaryons (anti-matter) were a sine qua non for the early persistence of our universe. It is now believed that the possibility of mass was predicated upon the detachment of the particle called the Higgs boson and with that the associated Higgs field. The detachment of the Higgs boson, and thus of mass, then constitutes a horizon for all subsequent material detachments in our universe.

Physicists characterize the possibility *state space* of a simple system in terms of its "degrees of freedom." For example, a simple atom like hydrogen can respond to a perturbation (such as being hit by a photon) by moving in space along three axes (3), rotating (4), or elevating the energy level of an electron (5). It is thus accorded five degrees of freedom. A simple diatomic molecule, like H_2, can also vibrate along its common axis so adds an additional degree of freedom. *Detachment is always about the emergence of higher degrees of relative independence.*

We can already see a reciprocal relationship between detachment and integration. Two atoms each with five degrees of freedom, shed their independent degrees of freedom through integrating, by way of a covalent bond, to form a new entity, the diatomic molecule, with six degrees of freedom. The more degrees of freedom the greater the detachment. As our universe has evolved it has given rise to constituent entities with greater and greater abilities to buffer themselves against "ambient winds" (be that bombardment by radiation or predation by voracious carnivores). Following the same logic, a particle with rest mass that creates a well in space-time is more detached than a particle (like a photon) with no rest mass. A macromolecule, like a

protein-based enzyme, whose folding history affects its future actions is more detached than a simpler molecule whose structure is solely determined by thermodynamic necessity (and thus has no history). A major transition in detachment occurred when a system emerged that actively constituted its own boundary (Plessner's *positionality*) and actively sustained its ability to do so. We typically associate this level of detachment with what we recognize as "life".

All states of detachment are relative, none are absolute. Levels of detachment exist in nested hierarchies. When a new level of detachment emerges, such as the boundary constituting, self-sustaining system (a simple cell), it also creates a space for *downward detachments* that may be viewed as parasitic on the higher level of detachment upon which it depends. Viruses emerged as expressions of downward (parasitic) detachment. Parasites and their hosts, lower and higher levels of detachment, dialectically interact resulting in the transformations of each and the appearance of new capacities that neither side of the equation alone could have produced. What begins as an "arms race" between parasites and hosts has resulted in de novo resources that enable parasite and host to re-integrate and make the jump to a new higher level of detachment. Internal compartmentalization and thereby an increase of organizational complexity has been a response strategy of "eukaryotic cells" to parasitic challenge (Koonin 2016) but through symbiotic re-integration has provided the basis for eukaryotes to engage in exploratory processes that lead the way toward both de novo ontogenetic and phylogenetic adaptations. The capacity for active exploratory processes ratchets up the degrees of freedom and level of detachment of a eukaryotic entity by exponential measures. When "variation" (in the Darwinian sense) is no longer principally about the stochastic roll-of-the-dice of enol-keto tautomerism in DNA replication, but rather is facilitated by the active processes of a detached entity and instigated by sensitivity to contingent ambient conditions, when might we be warranted to say that normative-choice making has come into play?[7]

The stark inadequacy of a narrowly-survivalist, Neo-Darwinism was amply revealed by the genome-theoretic debunking of the "junk DNA" thesis, which itself had been a leading inspiration for the far more popularly influential doctrine of the "selfish gene." It has been long-since empirically well-established that the vast majority of human genomic DNA not only doesn't code for unique protein sequences, nor even for non-coding regulatory elements, but rather is highly repetitive and virus-like in its structure. But rather than being merely a genomic "free-rider" as the proponents of the "junk DNA" thesis proclaimed, the endogenous, and potentially transposable and self-replicating sequences, have been shown to be instrumental in processes of genetic

[7] Kirschner and Gerhart (2005), two leading contemporary cell and developmental biologists, updated our understanding of the cellular processes involved in generating novelty, including that of "exploratory processes" and referred to these as the basis of "facilitated variation." That these concepts, based upon unimpeachable empirical evidence, and which radically change the perception of the status of cells and organisms as agents in evolutionary change, haven't filtered into general understanding should well raise questions about the warrants for privileging the continued promulgation of outdated versions of Neo-Darwinism as being more rational than myth, ideology or creed. The proper response to the potential dangers of Creationist irrationalism (if such there is) is not the defensive (or offensive) petrification of "classic ideas" that have been made refractory to new findings, new insights and new theories.

segmental duplication that result in species specific, perhaps even species defining, gene families (Moss 2006). The dialectics of detachment suggest that transitions to higher order levels of detachment will attract new rapprochements between competing entities from lower and higher levels of detachment, and thus that downward detachments, seen dialectically, are means toward a variety of different possible sequelae.

When Hegel announced that Spirit is the truth of Nature, he didn't imagine a story being told at a cell and molecular level and yet he offered a perspective that we may find insight-inspiring to generalize upon. The preponderance of philosophical normativists have kept naturalism, one way or another, at arm's length. The very idea, however, of an expansive movement of detachment, indeed a dialectics of detachment, anchors an impetus towards an inevitability of normativity as a force of integration, well beyond the imagination of our speculation-allergic normativists. If the legacy of propositionally-delimited reflection on the nature of normativity inhibits our ability to fathom normativity outside of modern human practice, then perhaps nature is calling upon us to start thinking beyond the constraints of this legacy. The case of the hybrid hominin, I suggest, will provide an enabling pathway toward this end (Moss 2016a). That is also to say, in the spirit of a renewed philosophical anthropology, that when we have met the true *missing link*, she will be us.

2.2 The hybrid hominin

Remarkably, the great majority of both popular and disciplinary accounts of human nature, human evolution, cultural evolution and the like stand oblivious to two elephants in the room, the biological juvenilizing of the human organism, and the uniquely human orientation toward cooperativity and "we-ness." Whereas the former insight has its origins in common-sensical observations of the late eighteenth century which progressively gained more recondite confirmation over the centuries,[8] the latter is arguably the leading psychological discovery of only twenty-first century human sciences.[9] Despite this temporal disparity, I will argue that these insights reflect mutually implicative aspects of the hominin we became and continue to be. Hominin neotenous juvenilizing and the hominin orientation toward 'we-ness' will be part of a story about a transition to a new level of detachment that will offer an exemplar for normative integration at a pre-conceptual level with implications ranging both back into pre-hominin existence and forward into the symbolically structured human world as well.

Our two elephants are typically ignored because they fail to coincide with deeply ingrained assumptions about the human individual; the individual human

[8] This insight is best thought of as beginning with Gottfried Herder's 1772 award winning "Essay on the Origin of Language" (1986 [1772]) taking theoretical shape with late nineteenth-century theories of neoteny, heterochrony and juvenilization and reaching its philosophical pinnacle with Arnold Gehlen's masterwork *Der Mensch* (1940) translated into English as *Man: His Nature and Place in the World* (1988). For details of the late nineteenth-century biological theories see Gould 1977, a recent exposition of the human neoteny view can be found in Bromhall 2003.

[9] The work of Michael Tomasello and co-workers at the Max Planck Institute for Evolutionary Anthropology has been a game-changer insofar as establishing the inherent orientation of human infants toward cooperative interactions and the centrality of the human capacity for social understanding and the capacity for "we-mode" as the unique cognitive differentia that distinguishes humans from other higher primates. An easily approachable introduction to these findings can be found in Tomasello 2009 and further explored in Tomasello 2019.

body and the individual human mind. These assumptions have been secured by various traditions and practices not limited to methodological individualism, Neo-Darwinist reductionism, neo-classical economics, rational choice theory, liberal political theory, and so forth. More recent attempts at elaborating cultural evolution theory from a Neo-Darwinist perspective simply follow suit in treating norms as bits of information that competing, self-interested individuals choose to adopt or not for reasons of individual instrumental benefit (Moss 2016b). At its most basic level, human sociality as such, has typically been side-stepped as if it were a non-question. Detachment theory offers a very different account (Moss 2014).

For cognitive psychologist Merlin Donald, the "riddle" of *Homo erectus* was the provocation that led to a breakthrough in thinking about human sociality (Donald 1991). How was it possible for *Homo erectus*, a species that endured for over a million years, that domesticated fire and lived in permanent encampments, that produced tools including the Acheulean hand-axe that would have required training to produce, that in greatest likelihood engaged in organized big mammal hunts that would have required a division of labor and who managed to leave Africa and colonize all of the contiguous Euro-Asian land masses (adapting to highly disparate environments and biomes), to do all of this without the benefit of spoken language?

Homo erectus, evidence compels us to conclude, had to have established a fully normatively integrated form of life, and for all intents and purposes was the first form of life to have done so. The thesis being proposed is that the evolution of *Homo erectus* constituted a radical transition in levels of detachment and we will proceed to draw on the work of various investigators, including Donald, to support this claim. Should this claim be accepted as warranted its implications would include radically undercutting the entire legacy of methodological, ontological and epistemological individualism, making common cause with some like-minded contemporaries in the areas of social ontology and phenomenology and reconfiguring the proper understanding of human freedom and autonomy.

If paleontologists agree on anything, they have agreed that early hominin survival, with the loss of the arboreal cover, required a level and a form of social cohesion unprecedented amongst higher primates. Everything we know about *Homo erectus* supports the view that an entirely new form of expanded sociality was achieved. But how was this possible? As early as the eighteenth century, Enlightenment thinkers turned their sights onto the human organism and thought about its relationship to the human mind. The common observation was that as organisms, humans are under-specialized weaklings compared to our fellow great apes. It was Gottfried Herder however who grasped the significance of this in detachment theoretic terms (Herder 1986 [1772]). For Herder, the loss of physiological specialization constitutes a detachment from the beck and call of nature (or natural stimuli) and thereby the precondition for a new form of directed attention he called *Besonnenheit.* On the basis of this simple insight Herder was able to become the grandfather of cultural anthropology. In the absence of behavioral determinations governed by instinctive stimulus-response mechanisms, a new level of integration could take place (and

indeed had to!). The hominin group could emerge on the basis of unprecedented degrees of freedom to constitute a way of life not by way of inborn-fixed response patterns but by way of the contingent constitution of group norms. Herder referred to a normatively integrated form of life as a Folk (*Volk*). The practices of a folk are its means of expressively constituting its way of being. Herder was thereby the founder of the expressive-constitutive theory of language[10] and of the study of folk practices in general that anticipated the birth of cultural anthropology as a discipline. When Franz Boas founded cultural anthropology in the last decade of the nineteenth century, it was expressly understood as a further manifestation of the Herder-Humboldt Volksgeist tradition (Bunzel 1996).

The transition to a new level of detachment in which system integration has become fully normative must have involved both losses and gains of functions such as to result in a net increase in degrees of freedom albeit at the level of the Group. The widespread observation of the radical dependency of the human neonate and the comparative weaknesses of the human individual speak to the evident loss of function; but from a biological standpoint how could these have come about? Evolutionary theory has increasingly become aware of the role that heterochrony, or changes in developmental timing, plays in evolutionary transitions. In the 1920s, the Dutch anatomist Louis Bolk put forward the fetalization thesis suggesting that major human features came as a package through the evolutionary retention of much of the fetal phenotype of the ancestral ape (Gould 1977). This thesis was revived in the form of a contemporary popular science account by English zoologist, writer and filmmaker Bromhall (2003). The ensemble of characteristic human features that resemble those of the fetal but not mature chimpanzee includes the following: an upright head, largely hairless body, massive brain with bulbous skull, flattened face, short lower jaw, small teeth, everted lips, in the female the retention of the outer labia, hymen and frontal position of the vulva, in the male the lack of a fully protective foreskin, lack of a penis bone and lack of spines along the penis. The Belgian anatomist Jos Verhulst also claims that human lungs and heart resemble that of the infant ape (Verhulst 2003). Alongside the fetalization thesis is the observation that the growth pattern of the human neonate's brain for its first year follows that, not of a great ape neonate, but rather that of a fetal brain giving rise to the idea of the human (or hominin) *extra-uterine year*.

By the lights of the fetalization thesis, the comparative enlargement of the human brain came as part of a systematic developmental package and was not initially selected for enhanced brain power (as the individualist outlook has traditionally assumed). Consistent failure to correlate differences in human intelligence with differences in human brain size would lend some credence to this view. The transition to a fully normatively regulated form of life, a transition that I refer to as the "First Detachment" (of hominin evolution), as already suggested, would necessarily involve both losses and gains – losses with respect to specialized response mechanisms triggered by particular natural stimuli, and gains with respect to the wherewithal for social integration and coordination. The extra-uterine pattern of brain growth can be seen as serving both of

[10] Charles Taylor has written on this extensively over his career but most comprehensively in his recent (2016) *The Language Animal – The Full Shape of the Linguistic Capacity* (The Belknap Press of Harvard University Press).

these requirements simultaneously. By relocating formative stages of brain development to the post-natal environment, basic structural formations become re-situated in a socio-cultural context. This transition fits hand in glove with the seminal work of behavioral ecologist Hrdy (2009). Hrdy's work is crucial for understanding the affective basis of the emergence of the Hominin Group (or supergroup). Comparing relative frequencies of parturition, length of developmental dependency and relative energy cost, Hrdy concluded that it would have been impossible for hominin mothers to have raised their offspring alone. The support of extended caregivers, *allo-parents*, would have been obligately required. Subsequent studies on post-natal and child-rearing behavior comparing women from various tribal communities with that of great ape mothers confirmed that great ape mothers are far less willing to allow others to hold their infant, and for over greater period of time, than human mothers, and further that only in the case of human mothers is there ever a rejection of a newborn due to some imperfection. Hrdy refers to the affective transition that is introduced by the sharing of infant caregiving at the earliest of stages as the onset of *emotional modernity* and she likewise concurs that allo-parenting and emotional modernity (i.e., first detachment) begins with Homo erectus. The hominin infant, underdeveloped and under-specialized at birth, and raised by an extended community of caregivers, became the first primate to be affectively well-suited to be a highly integrated member of a social group. The under-specialized neonate not only lacks the innate obstacles to normatively structured, socio-cultural inclusion, but is also in dire need of compensation for what she lacks. The hominin/human infant, as has become well established, has an appetitive drive toward shared attention and cooperative involvement (Tomasello 2009) driven by the need for compensation. The normatively structured world of the Group is the compensation.

Donald (2001) has led the way in terms of emphasizing that hominin evolution has been about the evolution of sociality. The hominin group coheres on the basis of shared emotions, shared perceptions and shared practices. Brains have evolved in relation to the cultures of cognition of which they are part. Mindedness is not an individual phenomenon but a cultural one.

Minds and cultures, being two reciprocal and interdependent aspects of a single phenomenon, are subject to changes that may be looked from either or both bottom up or top-down directions. But contrary to the common assumptions emanating from the reflective mind of the philosopher, there is no reason to assume that the nature of consciousness has been any more static than the nature of cultures and this will point will be elaborated upon further below. Does the kind of consciousness that allowed for First Detachment and the emergence of a flexibly adaptive normative form of life require or presuppose fully and characteristically (as *we* know it to be) self-conscious individuals? I don't see why it must (and indeed this will become a hallmark of "Second Detachment"). But even if full-on self-consciousness is not necessary for the transition to a fully normative form of life neither is merely the loss of environmentally oriented specialization nor is the affective openness to others sufficient. The lifestyle of *Homo erectus* was sufficiently complex that there had to be some medium for shared understanding at the requisite level of complexity, and a medium with sufficient semantic degrees of freedom to allow for coordinated responsiveness to contingencies beyond that which merely emotional contagion could provide for. Donald approached this problem, taking a cue from Vygotsky, and contemplating what

possible game-changing innovation could lie within a plausible *zone of proximate evolution*.[11] Great apes, notably chimps, gorillas and bonobos, are seen to possess fairly sophisticated levels of social cognition, yet only within the context of an in-the-present, social episode. Just on the basis of a fully upright stature, early small-brained hominins (i.e., australopithecines) would have already been able to take full advantage of the upright body as a canvass for expressive gesture within the episode. Donald reasoned that in light of an already highly developed capacity for motor coordination, hominins would be within an evolutionary stone's throw of gaining a de novo capacity for enhanced volitional movement outside of the episode, that mimics movements that have recognizable meaning within an episode. He referred to this capacity as that of *autocuing* (Donald 1991*)*. The capacity to autocue sequences of movements would have allowed *Homo erectus* deliberately to redeploy sequences of movements that already have meaning, with communicative intent. Donald refers to this as mimesis and makes the case for how mimesis would have transformed the culture and the cognition of *Homo erectus*. For present purposes, we can delimit our attention to the way in which mimesis could transform the capacity for normative integration. What does it mean to refer to an entity, or a group of entities, as engaging in a practice? For an activity to be a practice is to ascribe a normative content to it. To engage in a practice is to conform to the right way of doing something. Producing the Acheulean hand-axe was a practice, organizing a division of labor for a hunt would have been a practice, and any ritual, such as a ritual dance, was a practice. Donald suggests that with autocuing *Homo erectus* gained the capacity for "kinematic imagination", by which he meant the ability to imagine one's body acting in social context *as if* from an outside perspective. This would seem to coincide well with Helmuth Plessner's concept of "ex-centric positionality" (Plessner 2019) and suggests that ex-centric positionality would have likely begun with *Homo erectus*. This is not to say that *Homo erectus* would have enjoyed a full sense of selfhood as we know it, but rather that the embodied foundation of selfhood would have been established and provided the basis upon which symbolically mediated structures could eventually be built. With autocuing, hominins could become *implicitly* accountable for their actions. The culture and cognition of mimesis created the fabric, for the first time, of a fully normatively integrated form of life. With the culture of mimesis, Hominins entered the realm of Geist, not because they were engaged in explicit self-to-self recognition but because the medium of social integration was that of relations of normative as opposed to physical causality. *This claim does constitute an explicit, if perhaps subtle, challenge to the orthodox Hegelian view.*

3 Normativity, pre-reflective plural-self awareness and cerebral asymmetry

In recent literature concerned with collective intentionality there is a growing debate around the idea of "plural pre-reflective self-awareness." While it is somewhat more widely accepted that there is a minimal pre-reflective me-ness that accompanies our on-

[11] Vygotsky analyzed stages and transitions in human development guided by the concept of a zone of proximal development that spoke to what new developmental capacity could plausibly be within reach of what was already present. Donald "exapted" this concept for the evolutionary analysis.

going experience (which has been understood by Dreyfus and others as Heidegger's view), there are many hesitations about the assertion of a pre-reflective *us-ness*. But what these discussions lack, is exactly the genealogical account that the present theory provides. In order to round out our account of that normatively structured form of life that begins with First Detachment we must draw upon the recent renewal of research on cerebral hemispheric asymmetry inspired by the brilliant work of Scottish neuroscientist McGilchrist (2009). The story he has to offer, and upon which I have drawn some further extrapolations, provides exactly the kind of alternative to the individualistic and cognitivist account of the place of the normative that has impeded naturalistic enterprises from the get-go.

Why have cerebral asymmetry to begin with? The right and left hemispheres perform distinctly different and in principle complementary (and yet also competitive) functions. The right hemisphere is oriented toward the big picture, the holistic context, whereby the left hemisphere is oriented toward a discrete focus. The right hemisphere is responsible for all forms of attention except focused attention, i.e., vigilance, sustained attention, alertness and most of divided attention. Where the left hemisphere is analytic and logical and means-ends oriented, the right hemisphere is both the source of one's emotional style and character and the place where the emotions of another can be interpreted and understood. The right hemisphere is oriented towards end-in-themselves. Where the left hemisphere can judge logical consistency, it can't judge and detect the soundness of premises, even when patently absurd. The right hemisphere can judge the soundness of premises but not necessarily the analytic validity of inferences. As the right hemisphere is context sensitive it can detect and understand a frameshift (a change in context) where the left hemisphere cannot. The right hemisphere provides the location of the body schema and so was presumably instrumental in the transition to autocuing and kinematic imagination. The left hemisphere is parasitic on the right hemisphere to the extent of drawing upon its content for its own form of focal analysis.

That the right hemisphere is heavier in social animals provides some clue as to its history and trajectory. What seems very likely is that the right hemisphere played a role, especially for animals subject to predation, in providing an around-the-clock vigilance of the general surroundings. A moment's reflection will confirm that a) such general ongoing vigilance would make perfect sense for a creature subject to predation, and b) general vigilance and focused attention are two very different lines of work and having them separated into different hemispheres makes perfect sense as well. When and where mammals become social, the vigilance would concern itself also with the social dynamics of its surrounding environment. The key move I want to make at this point is to suggest that the right hemisphere was poised to become a normativity detector with the transition of First Detachment. We can gain a valuable picture onto what this might look like by referring back to the thesis put forward by former Princeton psychologist Jaynes (1976) in his celebrated *The Origin of Consciousness in the Breakdown of the Bicameral Brain*. Jaynes argued that there was no indication of self-consciousness as recent as the Iliad and that individuals acted according to what they experienced as the dictates of the voices of the gods that were heard as auditory hallucination produced by the right hemisphere. Jaynes further supported this thesis by the findings of studies done in Wilder Penfield's laboratory, whereby volunteers were subject to electrical stimulation of their right hemisphere resulting in auditory experiences, some but not all of which were experienced as voices, but all of which were experienced as originating

from a source external to the subject. It would follow, and all the more so given what we now know about the right hemisphere, that the right hemisphere was active in interpreting the normative meaning of its environment which it then announced to the left hemisphere. Consciousness, for Jaynes, in the familiar sense, begins when the left hemisphere commences the narration of a story to oneself about oneself which is based on extrapolation. Whether Jaynes is correct in his reading of the Iliad and his dating of the emergence of self-consciousness to as recently as within 3000 years ago is not crucial for present purposes. What Jaynes depicts with respect to the right hemisphere, pre-self-consciously monitoring the normative indications of its social environment is precisely what would enable *Homo erectus*, in the absence of speech, to achieve and enact a normatively integrated form of life prior to language. Jaynes' conjectures about the right hemisphere are consistent with the conclusions of McGilchrist decades later, which are also further supported by a wealth of intervening data.

The proposed primordial dominance of the right hemisphere and its role in ongoing normative perception, along with the affective and cognitive resources previously discussed, go a long way towards providing a plausible account of a pre-linguistic group existence, but I will suggest that one more piece of this puzzle must be put in place. I suggest that for a fully normatively integrated form of life to have flourished, as *Homo erectus* did for something approaching a million years, that there had to be a primordial norm around which all else paid implicit obeisance to, one and only one universal norm, and this was the norm that held that the Group is the Good. It is this Ur-norm, this magnetic north, around which all the practices of the group cohered and towards which they were inflected. It is this Ur-norm which allows an affectively and cognitively competent form of life to be constituted by practices that can thereby cohere together. The thesis of the hybrid hominin is that while we are no longer only creatures of the First Detachment Group and are now left-hemisphere dominant, self-narrating, self-conscious individuals, that we are *also still* creatures of the Group, hence the hybridity. The path toward Second Detachment individuation can also be approached in terms of the integrative force of normativity. Language, the spoken capacity for which is nested in the left hemisphere, was driven by the integrative benefit of rendering the content of ritual into myth. Likewise, the integrative benefits of individual accountability were initially strengthened by the capacity for reflective self-identification and linguistic mastery of the system of personal pronouns. In detachment-theoretic terms however, human individuation is also a form of downward parasitic detachment, which dialectically speaking has both been instrumental in challenging the limits of the traditional Group in the name of greater, more universal goods, and yet has also served to subvert the normative coherence of any level of human association. A more adequate elaboration of the normative implications of a dialectics of human detachment however must await another day.

There are many implications of the hybrid thesis which pertain to the scope and presence of the "normative force". Schmid (2014), who has defended the idea of a pre-reflective *plural self-awareness,* albeit as yet with no thought to a genealogical account, has called to evidence the experience of members of a group cleaving toward a normative desire for agreement as an apparent end-in-itself. Schmid's observation only touches the tip of the iceberg. The legacy of First Detachment continues to orient us toward abstract norms, the diffracted rays of the lost primordial group, above and beyond accountabilities to merely actual people at particular times and places. Children

of immigrants, for example, whose linguistic exposure is almost entirely limited to their parents during their critical language acquisition phase still manage to become native speakers of the adopted language with no residual accent. These offspring unconsciously privilege their perception of the ambient norm. Even the most accomplished intellectuals will curiously feel themselves to be at a moral disadvantage if they are outnumbered two to one in an argument, despite fully knowing that the vast majority of the as yet uncommitted remaining 7 billion extant humans may well side with their position. Empirical studies in social psychology have shown that, for example, being informed about the facts of waste disposal, waste pollution and recycling has less impact on the typical recycling practices of an individual than finding out what their neighbors are doing (independent of who their neighbors happen to be) (Kesibir 2012). Every term I teach I will invariably encounter students whose internalized norm about not raising their voice will trump the pleadings of a slightly hearing-impaired professor to speak up. The same student who refuses to ask a question, or express a view, in propria persona, will eagerly enter into a protracted peroration just so long as it is understood as the representation of a group of three. One may want to argue that there is "safety in numbers" but I would suggest that this adage is an easy cover for what is at root about the psychology of the Group and not an individualistic expression of a rational choice (about the benefit of numbers). Less anecdotally, I would suggest that the highly influential, if pragmatically ambivalent, Foucauldian concept of Power could be better materialized and made serviceable for human, and even meta-human[12] benefit, if it was reconstructed as an expression of normative force and further contextualized within a dialectics of detachment. Again, a topic for forthcoming work.

4 Coda

Finally, to return to the original question of the very hard problem. Even if it is granted that the thesis of the hybrid hominin can yield new insights into how the force of normativity does and/or doesn't constitute the glue of higher order unities in hominin and human life, is there any hope for imagining that any of this analysis can be backtracked down into progressive strata of pre-hominin life, let alone even so far as the early stages of the universe? Does a dialectics of detachment hold any greater promise for taking us closer to a fuller vision of nature beyond that of just so many parts outside of parts? Detachment theory suggests that all unities are only relative, that the unity we have taken to be most unequivocal, the unit of the self, is in fact derivative of the prior unity of the group, as well as a narrative construct that may conceal as much as it self-articulates. Detachment theory identifies a drive toward higher degrees of independence but in an even more speculative vein could we not also propose a *pain of detachment* that perhaps even begins with the Big Bang and constitutes a drive toward compensation for the loss of prior unities? Compensation for detachment? In true dialectical fashion, might not every transition in

[12] Post-humanism uncritically takes on board an individualistic misconception of the human and proceeds to build its worldview on the basis of an abstract negation of something it misunderstands. Transhumanism begins with the same misconception and moves in another misguided direction (see Moss 2017). Meta-humanism, building on detachment theory, will offer an alternative route for sublating anthropocentrism, not by eschewing the human but by exposing the expansiveness immanent of the hybrid hominin's normatively structured dialectics of detachment and compensation. We have become modern, but we were never *Human*.

detachment, constitute a compensatory re-integration on the one side and yet the pain-inducing provocation for further compensation on the other? Might it thereby be the case that if we were to at least loosen the ties of our neo-Darwinian and Newtonian vestments, that the hard problem of interiority (consciousness) and the *very* hard problem of integration might become indistinguishable? Could this be at least a glimpse of that truth that spirit reveals about nature?

References

Brandom, R. (2019). *A spirit of trust: A reading of Hegel's phenomenology*. Cambridge: Belknap Press.

Bromhall, C. (2003). *The eternal child*. London: Ebury Press.

Bunzel, M. (1996). Franz Boas and the Humboldtian tradition. In G. W. Stocking Jr. (Ed.), *Volksgeist as method and ethic: Essays on Boasian ethnography and the German anthropological tradition*. Madison: University of Wisconsin Press.

Chalmers, D. (1995). Facing up to the problems of consciousness. *Journal of Consciousness Studies, 2*(3), 200–219.

Donald, M. (1991). *Origins of the modern mind*. Cambridge: Harvard University Press.

Donald, M. (2001). *A mind so rare: The evolution of human consciousness*. New York: W.W. Norton.

Gare, A. (2018). Natural philosophy and the sciences: challenging science's tunnel vision. *Philosophies, 3*(4), 33–62 https://www.mdpi.com/2409-9287/3/4/33.

Gehlen, A. (1988). *Man: His nature and place in the world*. New York: Columbia University Press.

Gibson, D.G., Glass, J.I., Lartigue, C., Noskov, V. N., Chuang, R.-Y., Algire, M. A., Benders, G. A., Montague, M. G., Ma, L., Moodie, M. M., Merryman, C., Vashee, S., Krishnakumar, R., Assad-Garcia, N., Andrews-Pfannkoch, C., Denisova, E. A., Young, L., Qi, Z.-Q., Segall-Shapiro, T. H., Calvey, C. H., Parmar, P. P., Hutchison, C. A., Smith, H. O., Venter, J. C. (et al.) (2010) "Creation of a bacterial cell controlled by a chemically synthesized genome" Science 329:(5987), 52–56.

Gould, S. J. (1977). *Ontogeny and phylogeny*. Cambridge: The Belkhap Press of Harvard University Press.

Habermas, J. (1992). Individuation through socialization: On George Herbert mead's theory of subjectivity. In: Habermas, J. *Postmetaphysical thinking: Philosophical essays* (trans. Mark Hohengarten) Cambridge: Polity Press.

Herder, J.G. (1986 [1772]). Essay on the Origin of Language. In: *On the origin of language two essays. Jean-Jacques Rousseau Johann Gottfried Herder* (trans J. Moran & A Gode). Chicago: University of Chicago Press.

Hrdy, S. (2009). *Mothers and others*. Cambridge: Harvard University Press.

Jaynes, J. (1976). *The origins of consciousness in the breakdown of the bicameral mind*. New York: Houghton Mifflin.

Kesibir, S. (2012). The superorganism account of human sociality: how and when human groups are like beehives. *Personality and Social Psychology Review, 16*, 233–261.

Kirschner, M. W., & Gerhart, J. C. (2005). *The plausibility of life: Resolving Darwin's dilemma*. New Haven: Yale University Press.

Koonin, E. V. (2016). Viruses and mobile elements as drivers of evolutionary transitions. *Philosophical Transactions Royal Society B, 371*, 20150442.

McGilchrist, I. (2009). *The master and his emissary: The divided brain and the making of the western world*. New Haven: Yale University Press.

Moss, L. (2006). Redundancy, plasticity, and detachment: The implications of comparative genomics for evolutionary thinking. *Philosophy of Science, 73*(5), 930–946.

Moss, L. (2014). Detachment and compensation. *Philosophy & Social Criticism, 40*(1), 91–105.

Moss, L. (2016a). The hybrid hominin: A renewed point of departure for philosophical anthropology. In P. Honenberger (Ed.), *Naturalism and philosophical anthropology: nature, life, and the human between transcendental and empirical perspectives*. New York: Palgrave Macmillan.

Moss, L. (2016b). Review: Tim Lewens, cultural evolution: conceptual challenges, Notre Dame Philosophical Reviews. https://ndpr.nd.edu/news/cultural-evolution-conceptual-challenges/. Accessed 4 May 2016.

Moss, L. (2017). New naturalism and critical theory. In A. Gare & W. Hudson (Eds.), *For a new naturalism.* New York: Telos Press.

Plessner, H. (2019). *Levels of organic life and the human: An introduction to philosophical anthropology* (trans. M Hyatt). New York: Fordham University Press.

Schmid, H. B. (2014). Plural self-awareness. *Phenomenology and the Cognitive Sciences, 13*, 7–24.

Tomasello, M. (2009). *Why we cooperate*. Cambridge: The MIT Press.

Tomasello, M. (2019). *Becoming human: A theory of ontogeny*. Cambridge: Belknap Press.

Unger, R. M., & Smolin, L. (2015). *The singular universe and the reality of time. A proposal in natural philosophy*. Cambridge: Cambridge University Press.

Verhulst, J. (2003). *Developmental dynamics in humans and other primate: Discovering evolutionary principles through comparative morphology*. Adonis Press.

religions

MDPI

Article

The Implicit as a Resource for Engaging Normativity in Religious Studies

Gary Slater

Andre Hall 211, St. Edward's University, Austin, TX 78704, USA; gslater@stedwards.edu; Tel.: +01-512-233-1644

Academic Editor: Bharat Ranganathan
Received: 31 August 2017; Accepted: 6 November 2017; Published: 19 November 2017

Abstract: This piece recommends the *implicit* as a resource for examining normativity within the study of religion. Attention to the implicit serves at least two purposes toward this end. First, it gives the scholar of religion a clearer sense of the norms of the communities she seeks to understand, norms that, depending partly on one's methodological commitments, may be evaluated as well as described. Second, it deepens the scholar's reflections on the implicit norms that guide her own work. These claims—which extend the work of Tyler Roberts, Kevin Schilbrack, and Thomas A. Lewis—are embedded within specific understandings of language and mind as drawn from Robert Brandom and Peter Ochs. Brandom and Ochs help speak to the questions of whether the academic study of a religious tradition can or should evaluate that tradition, answering "yes" and "it depends", respectively. This presents scholars of religion with both a challenge and an opportunity. The challenge is that religionists no longer have recourse to a strict distinction between fact and value. The opportunity is that, by linking implicit facts and values to explicit analysis and evaluation, scholarly investigations can be expanded in both descriptive and prescriptive contexts.

Keywords: religious studies; method; logic; Robert Brandom; Peter Ochs; Scriptural Reasoning; normativity

1. Introduction

"Both read the Bible day & night/But thou readst black where I read white." These words comprise the final lines of William Blake's "The Everlasting Gospel." Scholars of religion wishing to understand Blake's words might read them in any number of ways. For example, a biblical scholar might find a poetic analogy for Sitz im Leben, imagining Blake's words as extolling the virtues of identifying the social setting of a text. A historian of religion might place the poem alongside Blake's *The Marriage of Heaven and Hell*, asking what insights into religion in romanticism these texts might reveal. An ethicist might find grounds to criticize Blake's claim in that text that the road of excess leads to the palace of wisdom. For the present paper, these lines from Blake are highlighted for a different reason: raising the *implicit* as a resource for examining normativity within the study of religion.

For present purposes, *implicit* is defined as "suggesting or suggested by something present to conscious awareness", and *explicit* as "present to conscious awareness". *Normativity*, for its part, is defined as "the state of being normative", with "normative" understood as "related to an evaluative standard". As with the "white" that Blake reads, the implicit is omnipresent, both distinct and inseparable from the explicit text. That is, discussing something as implicit automatically makes it explicit, and explicit texts depend in their intelligibility upon implicit background assumptions and possible responses. Attention to the implicit serves at least two purposes with respect to normativity and religious studies. First, it gives the scholar of religion a clearer sense of the norms of the communities she seeks to understand, norms that, depending partly on one's methodological commitments, may be evaluated as well as described. Second, it deepens the scholar's reflections on the implicit norms that guide her own work within the academy.

These claims are embedded within specific understandings of language and mind as drawn from the work of Robert Brandom and Peter Ochs. Since looking at the latent assumptions or consequences of cultural phenomena is already a widespread practice within the study of religion—employed by Marxist or psychoanalytic approaches, among others—it is important to clarify what makes the Brandom-Ochs pairing unique. There are at least three factors. First, as part of their affinity with pragmatic philosophy, Brandom and Ochs share a common understanding of discursive thought as inherently linked to purpose, or to put it in slightly different terms, linguistic meaning as inherently linked to use. Such emphases regarding purpose and use render these thinkers' ideas well suited to examining normativity, not simply because normativity is understood to be inescapable, but because the presumption of normativity allows Brandom and Ochs to make fine-grained distinctions regarding epistemic access to normativity and the implicit.

Second, following from the previous point, Brandom and Ochs collectively offer some key distinctions that are potentially helpful to scholars of religion. First, they place implicit alongside explicit reasoning in such a way that the former both contextualizes the latter and expands possibilities for further inquiry. Second, they allow one to distinguish *within* the implicit a heuristic of two types: (1) background assumptions brought by an author to render the explicit text intelligible, *antecedent* to the explicit text; (2) intended patterns of reception through actions or changed habits of reasoning on the part of the hearers/readers, *subsequent* to the explicit text. The former govern the selection of how to present an explicit text, and the latter govern the interpretation of that explicit text, including how one might act in response.

Third, and most important, the Brandom-Ochs pairing speaks to the question of whether the academic study of a religious tradition can or should evaluate that tradition. On whether they can, the answer is "yes". On whether they should, the answer is "it depends". The norms that govern religious practices do differ from those of academic discourses, and there may well be cases in which it is neither necessary nor appropriate for the scholar to engage in explicit normative evaluation. Broad differences between these two cultures—the academy, religion—certainly exist, and that is fine. Academic research per se entails commitment to falsifiability, for instance, that is not required within religious communities. But more effective than determining large general categorical differences between religion and the academy, I think, is attention to particular inquiries within particular contexts. The Brandom-Ochs pairing suggests that attention to local context is the most effective way to evaluate the norms of the culture one is studying. By providing the scholar with more information about those local norms, the Brandom-Ochs method of making implicit normativity explicit is a helpful part of this task. This method also serves similar ends with respect to the scholar's *own* norms of inquiry, helping to determine whether a given mode of inquiry is appropriate to that particular religious community in that context.

There are two readerships that this paper serves. The first includes those who are interested in questions of normativity and method in religious studies. Figures who fit this description include Donald Wiebe and Russell T. McCutcheon, who argue for a relatively sharp divide between the study of religion and religion itself. Also included are Thomas A. Lewis, Kevin Schilbrack, and Tyler Roberts, each of whom finds some space for normativity within their work. These latter thinkers are thus open to the possibility of a closer proximity between religious studies and religion than are Wiebe or McCutcheon. In a broad sense, this paper can be seen as an extension of the Lewis-Schilbrack-Roberts position within this conversation. This is because the paper uncovers a specific way to demonstrate the truth of the claim that normativity is pervasive across both religious and academic cultures.

This demonstration proceeds by engaging two problematic disagreements, applying the Brandom-Ochs approach to making implicit reasoning explicit. The first disagreement concerns the status of homosexuality in the United Methodist Church. The key participants here are the Confessing Movement Conference, which opposes gay and lesbian inclusion in the church, and the Reconciling Ministries Network, which favors it. The second disagreement concerns the status of normativity in

the study of religion. The key participants here are Donald Wiebe, who denies the normativity of religious studies, and Russell T. McCutcheon, who argues that religious studies is normative.

These examples are selected for two reasons. First, problematic debates are particularly useful in illustrating the benefits of making implicit reasoning explicit, as these debates tend to feature participants whose arguments are shaped by divergent, unrecognized norms. Second, with respect to these particular disagreements, one—on Wiebe and McCutcheon—reflects the norms of the academy, whereas the other—on the status of homosexuality in the United Methodist Church—reflects the norms of a religious community. The application of a common logic to these two cultures is itself a statement on the lack of a sharp methodological divide between religion and the academy.

The second readership for this paper is broader than the first. It includes religious studies scholars across various subdisciplines who might be interested in this particular demonstration of how normativity might be engaged within their work. This readership is especially broad because the Brandom-Ochs approach is relevant regardless of whether a scholar considers her work primarily descriptive or prescriptive. For descriptivists, making the implicit norms of a given religious community explicit is simply more information available for scholarly description.[1] This does not require evaluation, as it is possible to understand the norms of a community without subscribing to those norms themselves. For prescriptivists, the option to engage in normative evaluation is available when appropriate.[2]

This last point raises the issue of the norms among and between academic disciplines in the study of religion. The Brandom-Ochs pairing entails that normativity per se is pervasive. This pairing also recognizes that different disciplines have different norms even as they share a common identity within the academy.[3] This is to say that discipline-specific norms seek distinct types of understanding.[4] When evaluating religion seems appropriate, it is possible to do so.[5] But it is not always called for. For example, an anthropologist trying to grasp the purpose of a pilgrimage to a temple or a scholar on the Upanishads seeking the ideal translation from the Sanskrit does not also have to state her norms. The norms can stay implicit. Because the Brandom-Ochs pairing provides a logic of making implicit reasoning explicit in a way that is indifferent to any *particular* norm (even as it presupposes normativity per se), the option of making implicit normativity explicit is always available. It is the task of this paper to show how this is the case.

The essay proceeds in the following steps. In section two, it examines types of objections to normativity in religious studies. In section three, it introduces a set of authors—Roberts, Schilbrack, and Lewis—whose work responds to these objections. In section four, it introduces Ochs and Brandom and identifies these thinkers as part of a common tradition. In section five, it examines in turn each of these two thinkers' contributions to the making implicit reasoning explicit. And in section six, it applies Ochs and Brandom with respect to two examples, one drawn from the academy, on Wiebe

---

1 Elizabeth Bucar's article, "Bodies at the Margins: The Case of Transsexuality in Catholic and Shia Ethics" (Bucar 2010) is arguably an example of implicit norms within religious communities treated in a descriptive rather than prescriptive way. For further examples, see the focus issue, edited by Maria Heim and Anne Monius in the *Journal of Religious Ethics*, on the anthropology of moral worlds (Heim and Monius 2014).

2 It is even possible not merely to evaluate, but actually to help the religious community one is engaging by suggesting relevant responses to bad norms with attention to alternative normative resources within the same community. Ochs's idea of reparative reasoning (about which more below) embodies this, and it is the embodiment of the caretaker position McCutcheon attacks.

3 That discipline-specific norms seek distinct types of understanding is the reason this paper begins by invoking biblical studies, history, and ethics as different ways of investigating William Blake's "The Everlasting Gospel".

4 To provide just one example, consider historiography, whose objects of investigation exist within the following parameters: (1) the human past; (2) unique, non-replicable events; (3) topics amenable to the logic of explanation, in that the scholar can establish a logical fit between an agent's beliefs and motives and his or her actions.

5 I am hesitant to preemptively determine what these cases would be, though some disciplines, like ethics, are certainly more predisposed toward evaluation than others, like philology. This is a point that is made by Thomas A. Lewis, who argues that normativity is most overt in certain subdisciplines, such as ethics, philosophy, and theology, but also that it is to be found in all subdisciplines.

and McCutcheon, the other from a religious context, on the debate over homosexuality and the United Methodist Church.

2. Objections to Normativity

The tendency to divide religious studies into descriptive and prescriptive approaches often expresses itself in skepticism about normativity within religious studies. This skepticism exists in at least two forms. One form of skepticism grounds itself in a view of scientific objectivity as the paradigm of the academic study of religion. For example, according to Donald Wiebe, the legitimacy of religious studies as an academic discipline depends on a rigid separation between scientific and normative commitments. Wiebe sharply distinguishes between the scientific approach he extolls, on one hand, and preferential ties to some religious or political perspective, on the other. As Wiebe puts it, "Religious and political goals . . . are replacing the scientific agenda of seeking disinterested knowledge about religion and religions" (Wiebe 2005, p. 8), the implication being that scientific goals are descriptive whereas political-religious goals are normative. As Kevin Schilbrack notes, Wiebe's position entails a further distinction "between teaching religion and teaching about religion" (Schilbrack 2014, p. 191), a distinction upon which, as Wiebe sees it, the academic study of religion was founded.

A second form of skepticism about normativity in religious studies objects to the orientation of normativity toward religion rather than the presence of normativity itself. This position holds that the study of religion *is* normative, but also that its norms require that scholars maintain a strict critical distance from their objects of study. A leading proponent of this view, Russell T. McCutcheon, claims that scholars of religion should be "critics" rather than "caretakers" of religion, which precludes scholars from evaluating religious practices or claims on their own terms (McCutcheon 2001). McCutcheon holds that religious studies scholars do not "study religion, the gods, or ultimate concerns", but rather "use a folk rubric, 'religion', as a theoretically grounded, taxonomic marker to isolate or demarcate a portion of the complex, observable behavior of biologically, socially, and historically situated human beings" (McCutcheon 2001, p. 11). For McCutcheon, evaluation of religious practices or beliefs by way of norms internal to religion jeopardizes the critical distance he considers appropriate for scholarship. As McCutcheon puts it, "After all, the premise that makes the human sciences possible in the first place is that human behaviors always originate from within, and derive their culturally embedded meanings from being constrained by historical (i.e., social, political, economic, biological, etc.) entanglements" (McCutcheon 2001, pp. 7–8). McCutcheon's position entails that the distance between the norms of the academy and those of religion requires excluding the latter from the academic study of religion.

3. Responding to Skepticism about Normativity

There are a number of worthwhile commentaries on how to respond to skepticism about normativity within religious studies. Recent examples include Tyler Roberts's *Encountering Religion: Responsibility and Criticism after Secularism*, Kevin Schilbrack's *Philosophy and the Study of Religion: A Manifesto*, and Thomas A. Lewis's *Why Philosophy Matters for the Study of Religion—and* Vice Versa. These works variously suggest that because philosophy of religion engages questions of truth alongside those of value or normativity within religious belief or practice, the discipline is well placed to reflect on the implications of distinguishing description from prescription too rigidly. Yet these authors also argue that in spite of its potential relevance, philosophy of religion is failing to live up to its promise, whether through a lack of clarity about its scope, methods, and aims, an excessive parochialism, or a preferential focus on religious belief at the expense of religious practice.

Roberts argues not only that normativity within religious studies is inescapable, but also that normative religious claims are to be engaged and even celebrated. Rather than insulating religious studies *from* religion, as Wiebe or McCutcheon might have it, Roberts holds that "we might learn something about critical thinking by breaking back through *to* religion" (Roberts 2013, p. 168). Failing to recognize the resources *within* religion to help religious studies scholars become aware

of the academy's own limitations or modes of response to violence is an ethical as well as an academic failure. For Roberts, "it would be a failure to attend to practices and ideas that may offer alternatives to dominating and destructive ideologies, whether religious or not, and it would be a failure to know religion in all its complexity and power" (Roberts 2013, p. 6). In arguing his case, Roberts calls upon scholars to problematize the distinction between "religious" and "secular" in a manner that moves past the necessity of having to be "for" or "against" either the religious or the secular (Roberts 2013, p. 237). Escaping this binary opens new forms of encounter between the scholar of religion and her object of study.

Lewis, for his part, argues that normativity within religious studies is unavoidable, and so the relevant question is whether a given set of arguments can be publicly justified and examined. In offering resources for the examination of arguments, philosophy of religion should "conceive of this process broadly enough to encompass a wide range of justificatory strategies" (Lewis 2015, p. 55). In traversing the range Lewis is calling for, a helpful measure with respect to the present paper is the degree to which reflection on normativity is explicit as opposed to implicit. As Lewis notes, some disciplines, like ethics or philosophy or religion, are "*more likely* to be reflecting explicitly on the justification for their normative claims", whereas such disciplines as history or social sciences "are more likely to focus their energy elsewhere" (Lewis 2015, p. 53). Lewis thus suggests that the difference among the various disciplines that study religion is not so much the presence of normativity per se, but rather the extent to which the normativity is explicit. On this telling, Lewis addresses the debate in terms that speak to the concerns of the skeptics of normativity while suggesting a constructive path for both expanding and clarifying the discipline.

As for Schilbrack, he examines the arguments of the skeptics of normativity while attempting to construct an expansive, irenic role for philosophy of religion. He frames his vision in terms of three distinct-but-compatible goals: "*describing* religious phenomena, which must be done in terms of the agents' understandings; *explaining* those phenomena in terms of their causes; and *evaluating* the reasons that are or can be given for them" (Schilbrack 2014, pp. 179–80). On Schilbrack's account of these three tasks, it is evaluation that typically comes under the harshest critiques. Yet if "evaluative approaches are not part of the academic study of religions, the result will not be that evaluations are not included in the field, but rather that the evaluations already present in religious phenomena will be presented uncritically" (Schilbrack 2014, p. 187). Schilbrack's understanding of evaluation resembles Lewis's points about normativity: it is something that is present whether or not it is explicit, and reflecting on its presence in a self-conscious way is something that can help scholars improve their practices. An upshot of Schilbrack's commentary is that some effort to defend and clarify the task of evaluation within religious studies is therefore needed.

Certain claims follow from these commentaries. First, normativity is pervasive. Second, normativity has the capacity to be self-consciously acknowledged. As Lewis puts it, "normativity should not be avoided but rather self-consciously acknowledged and defended", entailing a "willingness to submit all claims to scrutiny and questioning" (Lewis 2015, p. 8). Schilbrack puts the matter in similar terms in holding that "the criterion for what belongs in the academy is not whether one's inquiries are value-laden—they always will be—but whether those values are open to challenge and critique" (Schilbrack 2014, p. 192). What I take Schilbrack to mean here is that one cannot deny that one has norms, nor can one exclude from the academic study of religion the philosopher who wants to examine them. Third, assumptions—which I take to belong to the category of the implicit—constitute a legitimate, perhaps even vital, area of inquiry for philosophy of religion. Such assumptions include those that enter into even questions of description within religious studies, including "what is or is not real, what can or cannot be known, and what activities are or are not worth pursuing" (Schilbrack 2014, p. 201). In sum, the commentaries of Roberts, Lewis, and Schilbrack suggest the following. First, normativity should be accepted. Second, normative arguments can be publicly debated. And third, assumptions are a vital topic of scholarly attention. Taken together, these claims trace a line that points toward a specific contribution from philosophy of religion in responding

to problematic attempts to divide description and prescription too deeply within religious studies methodologies. It is a task of this essay to map out one such contribution from Brandom and Ochs.

4. Turning toward the Implicit: Brandom and Ochs

Unlike Roberts, Lewis, and Schilbrack, this essay draws specifically from Robert Brandom and Peter Ochs. Brandom provides the intellectual resources needed for distinguishing implicit and explicit while maintaining analytical precision. Ochs self-consciously employs the implicit/explicit distinction in a religious context. The fact that Ochs considers his work as self-consciously theological does not detract from its legitimacy.[6] On the contrary, to recognize within a self-consciously theological project patterns of logic that are similar to the work of Brandom is to affirm the possibility of trespassing the divide between religion and the academy.[7]

In explaining how Ochs and Brandom speak to the debate over normativity within religious studies, it is necessary to set up the following taxonomy. As noted in the introduction, *implicit* is defined as "suggesting or suggested by something present to conscious awareness", and *explicit* as "present to conscious awareness". Description is likewise understood as an account of what something *is* and prescription as what something *should be/do*. There is also a clear difference between an *account*, as in prescriptive versus descriptive accounts, and a *form of reasoning*, as in implicit versus explicit

---

[6] There is an irony here. Reading Brandom and Ochs together suggests that normativity is pervasive within discursive thought. Yet, through its emphasis on logic, this approach also involves analytical tools that are not normative—though they are vague unless specified in context. By being falsifiable and "neutral", logic actually speaks to some of the critical, non-normative criteria in studying religion Donald Wiebe is calling for. The Brandom-Ochs approach also accords to a degree with McCutcheon. For instance, if one wishes to argue that the norms of the academy entail that scholars should be critics and not caretakers, there is no reason why this should contradict the claims about implicit normativity shown here. There is, however, *also* no reason to presume that the objects of one's inquiries—practitioners of religion—are not capable of the same critical reflection.

[7] In spite of their different interests and contexts, Brandom and Ochs share a common association with the pragmatic philosophical tradition. To be clear, invoking pragmatism is not necessary to demonstrate the relevance of implicit reasoning with respect to normativity in religious studies. Yet it is nonetheless helpful to do so. There are two reasons for this. First, invoking pragmatism makes the Brandom-Ochs relationship more coherent in a way that illustrates the possibility and the rewards of trespassing the boundary between religion and the academy. This is in keeping with the following suggestion from Roberts:

> We should consider . . . whether as scholars of religion we might learn something valuable by treating certain religious discourses not only as *objects* of study but as potential methodological *resources for* the study of religion and for cultural criticism. (Roberts 2013, p. 20).

Second, examining the pragmatic tradition is an important part of defusing the objection, explored below, that Wittgenstein and Peirce—thinkers on whom Brandom and Ochs respectively draw—are too different for their intellectual posterity to be read alongside one another.

It is true that there are many different versions of pragmatism, and that Ochs and Brandom are undoubtedly divergent figures within that tradition. But both thinkers are self-conscious about pragmatism in their own work and as a tradition. For one, the work of both scholars entails certain commitments associated with pragmatism, including the responsibility of philosophy to social and communal context, as well as what Robert B. Talisse and Scott F. Aikin have called "the three general rubrics of . . . clarity, coordination, and correctness" (Talisse and Aikin 2011, p. 5). Both scholars also invoke a common basic definition for pragmatism. Note the following from Ochs:

> My thesis is that pragmatic definition is not a discrete act of judgment or classification, but a *performance of correcting other, inadequate definitions of imprecise things*. Pragmatic reasoning is thus a different sort of reasoning than the kind employed in defining things precisely. It is a corrective activity. (Ochs 1998, pp. 4–5).

And note the following from Brandom:

> The more specific strategy by which the classical American pragmatists sought to naturalize the concept of *experience* . . . is what I will call *fundamental* pragmatism. This is the idea that one should understand knowing *that* as a kind of knowing *how* . . . That is, believing *that* things are thus-and-so is to be understood in terms of the practical abilities to *do* something. (Brandom 2011, p. 9).

Finally, both scholars are interested in pragmatism as a tradition, which suggests that my own invocation of pragmatism is at least partially consistent with the scholarly interests of the authors themselves. Brandom is the author of *Perspectives on Pragmatism*, and Ochs the author of *Peirce, Pragmatism, and the Logic of Scripture*, which understands the tradition in relation to its founding figure.

forms of reasoning. Accounts are explicit. Scholars of religion are in the business of generating accounts, and since these accounts are presented as papers, articles, or books, they are clearly explicit. For any explicit account generated, there are multiple implicit forms of reasoning, and these traverse the descriptive/prescriptive divide. That is, whether an account—inherently explicit—is prescriptive or descriptive, there are implicit forms of reasoning. These implicit forms of reasoning are both factual—in that it is factual that the forms of reasoning are implied—and value-laden, and they are accessible to analysis, at which point they have become explicit. The explicit and implicit are not binary opposites so much as they tend toward complementarity.

In exploring these distinctions with respect to Brandom and Ochs, there are two necessary caveats. First, this paper does not set out primarily to disprove or debunk those who wish to impose boundaries between descriptive and prescriptive accounts. Rather, its prime target is the over-extension of the prescriptive/descriptive distinction as a dualism or as a metaphysical dichotomy. On this point, a helpful definition can be found in Brandom, who holds that "a distinction becomes a dualism when its components are distinguished in terms that make their characteristic relations to one another ultimately unintelligible" (Brandom 1994, p. 615). Hilary Putnam has made a similar claim, arguing that "one difference between an ordinary distinction and a metaphysical dichotomy" is that "ordinary distinctions have ranges of application, and we are not surprised if they do not always apply" (Putnam 2002, p. 11). Putnam argues that the separation of facts and values "is, at bottom, not a *distinction* but a *thesis*" (Putnam 2002, p. 19).

The second caveat is that Ochs and Brandom do diverge on an important point: their intellectual inheritance regarding theories of meaning and reference. Brandom draws more from Ludwig Wittgenstein whereas Ochs draws more from C.S. Peirce. Although Wittgenstein and Peirce have left divergent legacies within contemporary philosophy, these thinkers do share a common affinity for rejecting the distinction between prescription and description.[8] Indeed, from Wittgenstein's critique of descriptivism in the *Philosophical Investigations*, which sought to shift philosophical attention to concrete contexts of language use, to Peirce's categorical equation of intelligibility with purpose, both thinkers offer strong arguments as to why a sharp dualism between description and prescription is simply untenable. When taken alongside Ochs's and Brandom's common appreciation for implicit/explicit as a methodologically useful distinction, the basic commonality between Peirce and Wittgenstein with respect to prescription/description warrants their enlistment toward a common purpose.

---

[8] Wittgenstein and Peirce has each left an enormous legacy regarding the relationship between meaning and reference, with Peirce credited as an independent cofounder of modern semiotics (along with Ferdinand de Saussure) and Wittgenstein possessing a vast legacy in the fields of logic, philosophy of mind, and philosophy of language, among other disciplines. Yet Wittgenstein and Peirce differ in that Peirce sought to build a philosophical architectonic to encompass all forms of inquiry within a common framework, dismissing efforts to delineate in advance what can and cannot be investigated, whereas Wittgenstein offered a theory of reference that explored "family resemblances", in which members of a given set share in certain overlapping traits rather than any single common quality that can be easily classified within a larger taxonomy. Wittgenstein and Peirce, in other words, are not philosophical allies.

Yet with respect to the present effort to read Ochs and Brandom together on implicit reasoning, they do not have to be allies. The simple reason for this is that, where it counts most, Brandom reads Wittgenstein in a way that is in harmony with the pragmatic tradition. For example, in his book *Between Saying and Doing* (Brandom 2008), Brandom argues that pragmatism remains most relevant if we understand pragmatics as providing special resources for extending and expanding the semantic analysis from concern with relations among meanings to encompass relations between meaning and use. Brandom argues for a methodological pluralism and integrative approach to discourse that is in fact more consistent with Ochs (and Peirce) than it is with Wittgenstein. As he puts it: "Rejecting scientism of the methodological monistic sort does not entail giving up the possibility of systematic philosophical theorizing about discursive practice" (Brandom 2008, p. 210). Instead of abandoning wholesale any sense in which language can refer beyond itself to an objective world, Brandom criticizes a more specific target: monistic exclusivism. As he puts it, "what is objectionable about the methodologically monistic form of scientism is its exclusivity" (Brandom 2008, p. 211). Because of these moves, Brandom and Ochs thus appear to be closer to each other on meaning and reference than their forebears in Peirce and Wittgenstein.

5. Robert Brandom's Analytic Pragmatism

To take each thinker in turn, Robert Brandom's work is instructive for offering a constructive vision for how the implicit/explicit distinction can circumvent the problems of too rigidly separating prescription from description. In *Making It Explicit*, Brandom traces the social conditions of discursive practices in a way that unites semantics and pragmatics. Brandom's dual interests in exploring normativity and the process by which implicit reasoning can be made explicit are of particular interest. As Brandom puts it in the book's preface:

> The practices that confer propositional and other sorts of conceptual content implicitly contain norms concerning how it is *correct* to use expressions, under which circumstances it is *appropriate* to perform various speech acts, and what the *appropriate* consequences of such performances are (Brandom 1994, p. xiii).

For Brandom, making something explicit is at the heart of discursivity as expression:

> One of the central tenets of the account of linguistic practice put forward here is that the characteristic *authority* on which the role of assertions in communication depends is intelligible only against the background of a correlative *responsibility* to vindicate one's entitlement to the commitments (assertional, inferential, and referential) implicit in an idiom without gainsaying the possibility of entitlement to a different one (Brandom 1994, p. xii).

These are technical passages that need unpacking. First, Brandom sees normativity as suffused throughout the process of interpretation, encompassing authority, responsibility, and entitlement. Second, in the relationship among authority, responsibility, and entitlement, it is the *implicit* that grounds responsibility and authority. In other words, normativity and the implicit are linked. As Brandom put it in his book *Perspectives on Pragmatism*, "norms that are explicit in the form of rules are intelligible only against a background of norms that are implicit in practices" (Brandom 2011, p. 69). Third, Brandom's claim that entitlement to implicit commitments need not gainsay those of other commitments suggests an appreciation of the logic of vagueness. This point about vagueness is exciting enough to merit further elaboration.

At the risk of overstating the matter, I see methodological richness in Brandom's implicit recognition of vagueness. Although vagueness is typically understood pejoratively as suggesting a lack of clarity, it is possible to define vagueness, logically, as form in which a given vague term has the capacity to be specified in an indefinite number of ways without exhausting its meaning. The significance of vagueness is something that has been recognized within philosophy of religion, particularly among thinkers influenced by C.S. Peirce.[9] As Peirce argued, in the logic of vagueness, the law of non-contradiction—a law of classical logic that holds that two contradictory propositions cannot both be true at the same time—does not apply (Peirce 1935, p. 355). As Peirce also put it, "A sign is objectively *vague*, in so far as, leaving its interpretation more or less indeterminate, it reserves for some other possible sign or experience the function of completing the determination" (Peirce 1935, p. 355). With respect to the relationship between vagueness and implicit reasoning, there are at least two possible permutations. A proposition can be vague in an explicit way—e.g., a fortune cookie that reads, "Your life is about to change dramatically"—in which case the specification remains implicit in the form of contradictory possible forms for dramatic life-change in the future. A proposition can also be *specific* in an explicit way, yet itself be the specification of a vague, implicit background assumption.

In the case of Brandom's reference to commitments "implicit in an idiom without gainsaying the possibility of entitlement to a different one", Brandom speaks to vagueness in a manner that bears on the ability to examine questions of justification, speaking therefore to the criterion set up by

---

[9] Examples of scholars of religion influenced by Peirce concerning the logic of vagueness include Ochs, Mark Randall James, Michael Raposa, and Robert C. Neville.

Schilbrack and especially Lewis. In particular, Brandom's reference to multiple implicit commitments entails an epistemology of vagueness that is both precise and pluralistic. For if the act of making reasoning explicit can also be an act of specifying the vague, and if, within vague reasoning, the law of non-contradiction does not apply, then one can understand scholarly accounts in religious studies as viably trespassing across the prescriptive/descriptive distinction at little cost to clarity.

Explicit accounts, by contrast, are specific and subject to the law of non-contradiction. This is essential to the process of description in religious studies—e.g., I am studying *this* religious practice, not *that* one, or *this* set of beliefs, not *those*. And yet since such accounts are understood to be explicit specifications of forms of reasoning that are implicit, vague, and ultimately normative—for the scholars themselves, as well as for their objects of study—then scholars also possess an opportunity to examine explicit accounts with respect to an inexhaustible multiplicity of normative frameworks, doing so without preemptively gainsaying any particular set of alternatives.

If the phrase "inexhaustible multiplicity of normative frameworks" appears to inject an unsavory element of arbitrariness into the process of inquiry, it bears pointing out that, in some ways, vagueness actually *restricts* the freedom of the scholar. As Mark Randall James put it in reference to Ochs, "a vague sign *restricts* the interpreter's freedom because even if it is true, a vague sign does not determine the consequences of its truth sufficiently for action in particular circumstances" (James 2016, p. 3). What is important here is that vagueness leaves a kind of *normed* freedom to the interpreter—she is free to wait, or inquire, but to do so subject to norms related to the object of inquiry. This speaks to the sense in which implicit normativity can be investigated in spite of its being ubiquitous.

Another helpful distinction that Brandom offers is that of the implicit as either *antecedent* or *subsequent* to the explicit. Brandom refers to the terms of this distinction as "language entry", which is antecedent to expression and perceptual, and "language exit", which is subsequent to expression and active (Brandom 1994, pp. 335–36). Brandom's distinction between antecedents and subsequents is helpful in that, though both antecedents and subsequents are implicit from the point of view of the explicit proposition, the former informs the effort of finding out the implicit thoughts behind an explicit text, whereas the latter bears on the implicit actions entailed by an explicit text.

Regarding logical antecedents in particular, Brandom's taxonomy speaks to the concerns about epistemic justification raised by Lewis and Schilbrack. As Schilbrack has put it, "The truth of a claim is logically independent of its source. But the *justification* of a claim is not independent of its source" (Schilbrack 2014, p. 202). If "source" is understood as logically antecedent, the relevance of Brandom's work to broader questions about engaging normativity within religious studies is clear. This is because Brandom identifies logical antecedents with implicit assumptions that precede "speech entry". If "source" as Schilbrack understands it and antecedents as Brandom understands them are related, then Brandom's explanation of how to make implicit antecedent assumptions explicit bears on explicating the source of a claim. Having been made explicit, such a source can be interrogated as to whether (and, if so, how) it appropriately justifies the claim that has stemmed from it.

Brandom's distinction between antecedents and subsequents is consistent with the pragmatic tradition's methodological commitments. Indeed, one finds a claim similar to Brandom's in the work of C.S. Peirce. As Peirce once put it, "The elements of every concept enter into logical thought at the gate of perception and make their exit at the gate of purposive action; and whatever cannot show its passports at both those two gates is to be arrested as unauthorized by reason" (Peirce 1935, p. 134). One must take care, however, not to overdraw Brandom's distinction between speech entry and speech exit. It is not always necessary—or even possible—to distinguish whether the implicit normativity that informs an explicit expression is prior or subsequent to it. Yet Brandom's speech entry/exit distinction is nonetheless helpful for the possibilities it offers in separating multiple implicit norms of entitlement (for speech entry) from multiple forms of vindication (for speech exit). This distinction also provides a link to the investigation of causality that is indispensable to social scientific studies of religion. As Schilbrack notes, "as we philosophers of religion shift our attention from questions of truth to include those of justification (and from issues of perception that are central to an individualist epistemology

to the issues of trustworthiness and credibility that are central to a social epistemology), I judge that we will come to see the assessment of religious claims as necessarily in conversation with the causal explanations provided by the social sciences" (Schilbrack 2014, pp. 202–3). Among its other promising entailments, Brandom's work makes this link explicit.

6. Peter Ochs's Scriptural Reasoning

For its part, Peter Ochs's work instructively demonstrates how the implicit/explicit distinction can be applied with respect to religion. As noted, Ochs is a leading figure within Scriptural Reasoning, a postliberal interfaith dialogue movement encompassing Christian, Jewish, and Muslim texts. Even though Ochs tends to be reticent about extending his work beyond its hermeneutical base in scriptural communities, his method of what Nicholas Adams has called "reparative reasoning" (Adams 2008, pp. 447–57) is potentially amenable to applications beyond postliberal theology, including the study of religion. The term "reparative" in Ochs's work can be understood in relation to the disciplines in which Ochs operates. As Adams puts it:

> One of the most arresting fruits of Ochs' method arises from the conjunction of scientific inquiry, historical investigation and ethnographic description. Ochs *qua* scientific inquirer fashions hypotheses, in response to real doubts, until new beliefs are established and taken as axioms which guide habits. Ochs *qua* historian investigates his own rabbinic tradition's interpretations of scripture as pragmatic responses to real doubts, issuing in new beliefs taken as axioms which guide habits, within that tradition. Ochs *qua* ethnographer attends to others' practices, which he takes to be responses to real doubts, and reconstructs the ways in which those others establish new beliefs which are taken as axioms, within those other traditions, which guide habits. (Adams 2008, p. 450).

There are two areas that are most helpful within Ochs's work with regard to the implicit/explicit distinction and religious studies: (1) the sense in which the interpretation of religious texts is shaped by implicit habits of reasoning intelligible with respect to community-specific, temporally-thick trajectories of historical reception; (2) the capacity of logic to disclose contexts in which the inherent vagueness of language can be specified in experience.

Since the present appropriation of Ochs's project varies in its aims from those of Ochs himself, it is worth unpacking his method of inquiry as he himself characterizes it. In *Peirce, Pragmatism, and the Logic of Scripture* (Ochs 1998), Ochs describes the implicit/explicit distinction as a means to repair problems of interpretation for specific communities. The first step in his method—after resolving to respond to a specific problem—is to "collect explicit texts as collections of particular arguments", understanding that each argument entails a set of implicit arguments (Ochs 1998, p. 24). The presumption is that some of these implicit arguments must include logical contraries of another *explicit* argument, and vice versa. Such is the initial diagnosis of a problem. The next step in Ochs's method involves "raising various hypotheses about the references of the text", which suggests bringing implicit arguments to the surface as a means of tracing out the fissures—as well as the relevant readerships and purposes—for the arguments of the explicit text (Ochs 1998, p. 26). Having generated some possible candidates for the source of a problem from among the implicit arguments, Ochs's method proceeds to call for distinguishing "between the explicit . . . text and the implicit text of which it is a sign" (Ochs 1998, p. 32). The final step is "to explicate the implicit text, transforming it from indefinite sign of some problem in some world to a general sign that recommends to an interpreter methods of solving that problem" (Ochs 1998, p. 33). Following from these steps, all that remains is the performance, that is, the subsequent action on the part of an interpreter that enacts her commitment to interpreting a general sign in lived experience.

Ochs's method overlaps with that of Brandom, whose work likewise suggests the implicit as the domain of vagueness and normativity. Yet in Ochs's case the link between normativity and implicit reasoning is expressed in different terms. For example, Ochs characterizes Scriptural Reasoning as

operating by way of the dialectic between diagrammatic and corrective functions. To diagram is to locate within an interpretive framework. This is a move that entails both explanation and description. To correct is to respond to some problem that motivates an inquiry—and inquiry, for Ochs, is always a response to a problem—in a manner that prescribes a pathway for amelioration. Ochs holds that "the diagrammatic-and-corrective reading one uses to repair pragmatic writing is the same reading one uses to prove its validity" (Ochs 1998, p. 277). He also argues that one's inquiries, "from out of the vague continuum of common sense . . . reify selective legislations that address some particular dialogue between theoretical resources and practical needs" (Ochs 1998, p. 268).

Ochs's diagrammatic-corrective continuum maps onto Brandom's antecedent/subsequent distinction regarding implicit reasoning. The "vague continuum of common sense" can be understood as antecedent to an explicit iteration, whereas the performance that is the proof of a pragmatic reading can be understood as at least partially subsequent to it. To be sure, Ochs does discuss implicit reasoning in a different voice than Brandom. Ochs, for instance, highlights both implicit performances that follow from and implicit rules of reasoning that precede—and, in some sense, determine—explicit ones, whereas Brandom tends to discuss the implicit in the form of norms that govern practices (Brandom 2011, p. 69). Ochs's normative goals also differ from those of scholars from across most of the disciplines that comprise religious studies, whether these may be putatively descriptive (e.g., sociology, anthropology) or prescriptive (e.g., ethics, theology). Yet the notion of implicit forms of reasoning is capacious enough to include both implicit arguments entailed by explicit ones (Ochs) and also patterns of action in continuity with explicit expressions (Brandom). Both visions are anti-dualist, anti-foundationalist, and thoroughly pragmatic.

7. Two Examples

What do these observations entail for religious studies? For one thing, the implicit/explicit distinction provides a means to root out unexamined biases when they have become problematic, combatting what Jonathan E. Brockopp calls "incidental normativity", which "occurs when our frames of reference seem obvious, causing us to overlook alternative interpretive possibilities" (Brockopp 2016, p. 28). It is worth substantiating this point by offering a sketch of how the implicit/explicit distinction can facilitate inquiry. Consider two examples—one drawn from a religious context and the other from an academic one. Both examples are thoroughly normative. The first example concerns a recent debate over the place of homosexuality in the United Methodist Church. The issue of homosexuality in the church has divided Methodists since as far back as its General Conference of 1972, with recent decades witnessing the formation of such rival groups as the Confessing Movement, which opposes recognition of gays and lesbians in the church, and the Reconciling Ministries Network, which favors it. At the 2012 United Methodist General Conference in Tampa, reformers among both clergy and laity unsuccessfully lobbied to have the Church's Book of Discipline amended to remove statements that prohibit clergy from performing same-sex marriages and condemn homosexual acts.

In spite of professions of unity from both sides, Methodists might find themselves agreeing with Ethan C. Nobles that no compromise on this issue is ultimately available, and that United Methodists are headed for a split akin to that which affected the Presbyterian Church in 2011 (Nobles 2012). The following passage comes from the official declaration of the 2005 Confessing Movement Conference, which opposes accommodation of homosexuality in the Methodist Church:

> Genuine unity in the church is not secured by religious sentiment, sincere piety, tight property clauses, or appeals to institutional authority and loyalty . . . Genuine unity, as a precious gift of the Holy Spirit, is rooted in the gospel of Jesus Christ, witnessed to in the Holy Scripture, summarized in the ecumenical creeds, celebrated in worship and sacraments, demonstrated in common mission, articulated in our teaching, lived out in love, and contended for by the faithful. (Confessing Movement Conference 2005).

A contrasting view comes from the New England delegation to the 2012 General Conference, on behalf of the Reconciling Ministries Network. The Reconciling Ministries Network supports recognizing gay members of the church:

> We will not be saved by our bishops, our polity, our structure, our metrics, our theology, our doctrine, our social principles . . . Our strength and our unity lie in our identity as a spiritual movement, grounded in the grace of God and linked by common practices of personal and social holiness. Nothing more, nothing less. (New England Delegation to the United Methodist 2012).

In spite of the explicitly opposing views, these two statements share a common grammar, revealing several implicit commonalities. Both arguments employ implicit rules of reasoning that shape the respective arguments into binary oppositions, each of which hinges upon contrasting definitions of unity. In the case of the Confessing Movement, the binary embedded in this passage is between two definitions of unity, and the authors of this passage clearly intend for the reader to understand that unity based on "the gospel of Jesus Christ" and a "common mission" is superior to that based on "tight property clauses" or "appeals to institutional authority". In the case of the Reconciling Ministries Network, the binary also bears on the question of unity in the Church, offering a choice between unity based on ecclesiastical structures and unity based on "the grace of God" and "personal and social holiness". In the case of the shared implicit binary, this is "unity means favored position on homosexuality in the Church, disunity means opposing position on homosexuality in the Church". Neither argument explicitly recognizes more than one framework of interpretation by which the issue of homosexuality and the Methodist Church can be examined.

As a second example of facilitating inquiry through attention to implicit reasoning, take the implicit assumptions that shape the arguments of Donald Wiebe and Russell McCutcheon regarding the exclusion of at least certain forms of normativity from academic religious studies. Note the following statement from Wiebe:

> I see [science and religion] as divergent or incommensurable modes of thought. The scientist would not, that is, talk of holy things but only of the cultural postulation . . . And the scientist would, moreover, consider it entirely appropriate to put to scrutiny claims about, and to analyze behaviour in relation to, such culturally postulated realities. (Wiebe 1992, p. 66).

Here is a corresponding statement from McCutcheon:

> An apology for the study of religion in the modern academy that presumes scholars of religion to be empathetic caretakers and naïve, well meaning hermeneuts is doomed from the outset, for it fundamentally confuses a distinction that lies at the base of the human sciences, between theoretically based scholarship on assorted aspects of human behavior and those very behaviors themselves. (McCutcheon 2001, p. 17).

Although Wiebe and McCutcheon's perspectives differ from one another, the explicit texts here are clearly not antagonistic in anywhere near the same manner as in the debate over the United Methodist Church. Yet in logical terms, it is possible to observe a similar pattern at work at the level of implicit reasoning. As with the statements from the Confessing and Reconciling Movements, Wiebe and McCutcheon are expressing opposing sides of an implicit distinction held in common between the two. In Wiebe's case, the religious studies scholar is scientific, whereas for McCutcheon the scholar is governed by norms; yet for both views, the ability to evaluate and argue about religion is implicitly verboten. There is a further implication here, which is that since religion cannot be argued about in a scholarly way, then argument about religion belongs to faith, thus upholding a strict faith/reason divide—a point that Lewis has similarly noted (Lewis 2015, p. 45). For both thinkers, scholarly inquiry is either "scientific" (Wiebe) or it is normative (McCutcheon), with the secular academy supplying the appropriate methodological distance from religion in either case.

Having briefly sketched the implicit reasoning that informs and determines the explicit texts for these two examples, it is possible to develop an outline for how the implicit/explicit distinction might be useful to scholars of religion. There are three sequential steps in such an outline: identifying normativity within the explicit text, excavating implicit normativity as antecedent assumptions, and evaluating implicit normativity as subsequent actions or entailments:

(*1*) *Identifying normativity within the explicit text*

In the case of the debate over homosexuality in the United Method Church, the respective norms are grounded in the term "unity" to support opposing positions on inclusion. In the case of Wiebe, McCutcheon, and the exclusion of certain norms within academic study of religion, the normative thrust points toward scholars embracing scientific norms/critical norms for Wiebe and McCutcheon, respectively.

(2) *Excavating implicit normativity as antecedent assumptions*

The implicit as an antecedent category is intimately bound up with epistemic justification. And on that question, a common pattern applies for both examples: different antecedent assumptions can be traced toward opposing sides of a shared implicit binary distinction. In the Methodism example, these antecedent assumptions are "unity means choosing tradition over contemporary culture in remaining opposed to homosexual congregants" and "unity means choosing contemporary culture over tradition in welcoming homosexual congregants" on the parts of the Confessing Movement and Reconciling Ministries, respectively. In this case, culture/tradition is a shared implicit binary, in that both sides assume that key aspects of United Methodist tradition are incompatible with contemporary culture on the issue of homosexuality. In the Wiebe-McCutcheon example, these antecedent assumptions are "religious studies as scientific precludes the scholar from prescriptive analysis of religious belief and practice" and "religious studies as normative within the modern academy precludes the scholar from prescriptive analysis of religious belief and practice" on the parts of Wiebe and McCutcheon, respectively. In this case, prescriptive analysis of religion/scholarship is a shared implicit binary.

(*3*) *Evaluating implicit normativity as subsequent actions or entailments*

For Wiebe and McCutcheon, as well as for the participants in the Methodist debate, the implicit entailments of the explicit views are exclusionary. Were one to act on the conclusions that McCutcheon, Wiebe, or the participants in the Methodist debate prescribe, the result would invariably be the exclusion of some subset of a common normative community—whether this community is the academy or the Methodist Church. In these examples, exclusion is by no means merely implicit; in its own distinctive way, each of the above arguments is *explicitly* exclusionary, as well. Yet in each case, such exclusion runs contrary to other aspects within the norms of that community. For example, the norms of welcoming the stranger and loving one's enemies are particularly claimed by the Methodist tradition. With religious studies, the humanistic ethos of "nothing human is foreign to me" reflects a particular normative commitment. To evaluate implicit normativity as subsequent action is thus also to speculate on the impact of the explicit claims with respect to the respective larger normative communities implicated. As Ochs's continuity between diagrammatic and corrective reasoning suggests, such evaluation is also the first step in a constructive response.

8. Conclusions

Brandom's distinction between implicit and explicit and Ochs's critical investigations of binaries can be combined in a highly compelling way. The result presents scholars of religion with a challenge and an opportunity. The challenge comes because undermining a rigid prescription/description distinction likewise undermines a sense in which scholars possess a surefire means to separate facts from values as they investigate the truth of religious claims. The opportunity comes because values are

no longer locked within one's subjective experience, and are potentially open to logical investigation in a way that would have previously been thought impossible. As one inquires into the social practices that attend the passage between the (explicit) description of some observed phenomenon and the (implicit) assessment that deems the description warranted, one finds a fusion between two forms of prescriptive thought: how to think about the phenomenon and what to do in response to it. Likewise, a given normative measure, though possessed of its own norms and responsible to the objects within its contexts, is part of a larger constellation of norms that one may attempt to make explicit—or not—as the interpretive situation demands. In so doing, erstwhile habitual, implicit, and general orders become rendered explicit and brought into conscious analysis.

Conflicts of Interest: The author declares no conflicts of interest.

References

Adams, Nicholas. 2008. Reparative Reasoning. *Modern Theology* 24: 447–57. [CrossRef]

Brandom, Robert. 1994. *Making It Explicit: Reasoning, Representing, & Discursive Commitment*. Cambridge: Harvard University Press.

Brandom, Robert. 2008. *Between Saying and Doing: Towards an Analytic Pragmatism*. Oxford: Oxford University Press.

Brandom, Robert. 2011. *Perspectives on Pragmatism: Classical, Recent, and Contemporary*. Cambridge: Harvard University Press.

Brockopp, Jonathan E. 2016. Islamic Origins and Incidental Normativity. *Journal of the American Academy of Religion* 84: 28–43. [CrossRef]

Bucar, Elizabeth. 2010. Bodies at the Margins: The Case of Transsexuality in Catholic and Shia Ethics. *Journal of Religious Ethics* 38: 601–15. [CrossRef]

Confessing Movement Conference. 2005. Unity in Christ, that the World May Believe. The Confessing Movement within the United Methodist Church. Available online: http://confessingumc.org/our-story/2005-unity-statement/ (accessed on 7 October 2016).

Heim, Maria, and Anne Monius. 2014. Recent Work in Moral Anthropology. *Journal of Religious Ethics* 42: 385–92. [CrossRef]

James, Mark Randall. 2016. Post-Liberalism as Wisdom: Ochs on Vagueness and Inquiry. Paper presented at the 2016 AAR/SBL Annual Meeting, San Antonio, TX, USA, November 20.

Lewis, Thomas A. 2015. *Why Philosophy Matters for the Study of Religion—and Vice Versa*. Oxford: Oxford University Press.

McCutcheon, Russell T. 2001. *Critics Not Caretakers*. New York: SUNY Press.

New England Delegation to the United Methodist. 2012. New England delegation: UMC's Identity and Unity Must Be Spiritual, Not Structural. *The United Methodist Reporter*. Available online: http://www.unitedmethodistreporter.com/2012/06/new-england-conference-delegation-no-common-identity-for-umc/ (accessed on 7 October 2016).

Nobles, Ethan C. 2012. Between a rock and a hard place.... *First Arkansas News*. Available online: http://firstarkansasnews.net/2012/05/between-a-rock-and-a-hard-place/ (accessed on 6 October 2016).

Ochs, Peter. 1998. *Peirce, Pragmatism, and the Logic of Scripture*. Cambridge: Cambridge University Press.

Peirce, Charles Sanders. 1935. *Collected Papers of Charles Sanders Peirce, Volumes V and VI: Pragmatism and Pragmaticism and Scientific Metaphysics*. Edited by Charles Hartshorne and Paul Weiss. Cambridge: Harvard University Press.

Putnam, Hilary. 2002. *The Collapse of the Fact/Value Dichotomy and Other Essays*. Cambridge: Harvard University Press.

Roberts, Tyler T. 2013. *Encountering Religion: Responsibility and Criticism after Secularism*. New York: Columbia University Press.

Schilbrack, Kevin. 2014. *Philosophy and the Study of Religions: A Manifesto*. Hoboken: Wiley-Blackwell.

Talisse, Robert B., and Scott F. Aikin. 2011. Introduction. In *The Pragmatism Reader*. Edited by Robert B. Talisse and Scott F. Aikin. Princeton: Princeton University Press.

Wiebe, Donald. 1992. On the Transformation of 'Belief' and the Domestication of 'Faith' in the Academic Study of Religion. *Method and Theory in the Study of Religion* 4: 47–67. [CrossRef]
Wiebe, Donald. 2005. The Politics of Wishful Thinking? Disentangling the Role of the Scholar-Scientist from that of the Public Intellectual in the Modern Academic Study of Religion. *Temenos* 41: 7–38.

doi:10.1017/aju.2020.17

SYMPOSIUM ON PROSPER WEIL, "TOWARDS RELATIVE NORMATIVITY IN INTERNATIONAL LAW?"

THE INTERNATIONAL LAW OF CO-PROGRESSIVENESS AS A RESPONSE TO THE PROBLEMS ASSOCIATED WITH "RELATIVE NORMATIVITY"

Sienho Yee[*]

Prosper Weil misfired his volleys by targeting his protestations at relative normativity in international law. In itself, relative normativity is unavoidable and beneficial. The enduring value of his celebrated 1983 article "Towards Relative Normativity in International Law?"[1] lies in identifying the various problems that he associated with relative normativity. These problems deserve serious attention and conscientious responses in order to assure the health of the international legal system as well as the international community. The idea of an international law of co-progressiveness that I have developed,[2] though not intended as a direct response to these problems, does come with a toolkit full of responses that would go a long way to solving those problems or at least reducing them to a minimum.

Problems Associated with Relative Normativity

Weil detailed many problems associated with the relativization of norms; to varying degrees, these problems challenge what he viewed as the three pillars of the international legal system: voluntarism, positivism, and neutrality. However, the problems he identified are not necessarily a result of the relativization of norms, but may be a result of other forces, or are even structurally inherent in the international legal system. Indeed, Weil's article's lasting value lies in identifying and propounding these problems, to the point of exaggeration. Here I will highlight some, but not all, of them.

One type of problem goes to imperfect participation in law-making, especially regarding *jus cogens*, essential norms, obligations *erga omnes*, obligations *omnium*, and other super-norms. In the formation of such norms, Weil notes,

> a rule acquires superior normative density once its preeminence is accepted and recognized by "all the essential components of the international community." But since a state's membership in this club of

* *Changjiang Xuezhe Professor of International Law, China Foreign Affairs University, Beijing; visiting professor, Faculty of Law, University of Macau; Chief Expert, Wuhan University Institute of Boundary and Ocean Studies.*

[1] Prosper Weil, *Towards Relative Normativity in International Law?*, 77 AJIL 413 (1983).

[2] Sienho Yee, *Towards an International Law of Co-Progressiveness*, *in* INTERNATIONAL LAW IN THE POST-COLD WAR WORLD: ESSAYS IN MEMORY OF LI HAOPEI 18–39 (Sienho Yee & Wang Tieya eds., 2001) (reprinted in SIENHO YEE, TOWARDS AN INTERNATIONAL LAW OF CO-PROGRESSIVENESS 1–26 (2004)); Sienho Yee, *The International Law of Co-Progressiveness: The Descriptive Observation, the Normative Position and Some Core Principles*, 13 CHINESE J. INT'L L. 485–500 (2014); SIENHO YEE, TOWARDS AN INTERNATIONAL LAW OF CO-PROGRESSIVENESS, PART II: MEMBERSHIP, LEADERSHIP AND RESPONSIBILITY (2014) [hereinafter YEE, PART II].

> "essential components" is not made conspicuous by any particular distinguishing marks—be they geographical, ideological, economic, or whatever—what must happen in the end is that a number of states (not necessarily in the majority) will usurp an exclusive right of membership and bar entry to the others, who will find themselves not only blackballed but forced to accept the supernormativity of rules they were perhaps not even prepared to recognize as ordinary norms.[3]

Second, to Weil, alongside or associated with the problematic participation in lawmaking is the problematic expansion of the scope of application of international law, from the original bilateral or conventional scope, to the general scope (via customary or general international law), then to the universal scope (via "universal law"). He took as signs of such danger the International Court of Justice's analysis of the relationship between treaties and custom in *North Sea Continental Shelf*,[4] and its penchant in *Hostages in Tehran* for referring to "obligations under general international law" despite the apparent sufficiency of the convention at issue in grounding the case.[5] The result of such an expansion is that, in *some* situations, for a dissenting state there is not only no chance to participate in the formulation of the norm, but also *no exit* from the application of that norm, because the escape hatch—the persistent objector rule—is not available. Here Weil observed a transition from the classic "presumptive acceptance to imposed acceptance."[6] The circle of imposition is thus complete. This phenomenon is now reflected in the International Law Commission's Draft Conclusions on *jus cogens*, adopted on first reading, Conclusion 14(3) of which states that "[t]he persistent objector rule does not apply to peremptory norms of general international law (*jus cogens*)."[7]

Third, concomitant with the expansion of the scope of application of international law is the expansion of the interest of a state in the enforcement of international law, or the standing of a state in that regard, which Weil found to be alarming. For him, the international legal system is such that "it is up to each state to protect its own rights; it is up to none to champion the rights of others."[8] The arrival of the concepts of obligations *erga omnes*, obligations *omnium*, etc., opens the door for the idea of *actio popularis*. Left unregulated,

> that would mean that any state, in the name of higher values as determined by itself, could appoint itself the avenger of the international community. Thus, under the banner of law, chaos and violence would come to reign among states, and international law would turn on and rend itself with the loftiest of intentions.[9]

Fourth, not only were the movement from non-norm to super-norm, the proliferation of super-norms, and the expansion of the scope of application concerning to Weil, but so also was the excessive speed of all that was happening. For him, the "stealthy rise" of non-norm to super-norm was already a process difficult to contain.[10] He saw this rush as undesirable: "By seeking to create today the law of tomorrow's international society, one runs the risk of cutting a key that will not fit the lock it will have to open."[11]

Fifth, all these problems would finally result in great uncertainties in the law, Weil feared. Indeed, he invoked "uncertainty" six times (in the singular or in the plural) in his article to describe the new situation. To take just one

[3] Weil, *supra* note 1, at 427.

[4] *Id.* at 436-37.

[5] *Id.* at 439.

[6] *Id.* at 437.

[7] Int'l Law Comm'n, Report on the Work of Its Seventy-First Session, UN Doc. A/74/10, at 182 (Apr. 29-June 7, July 8-Aug. 9, 2019).

[8] Weil, *supra* note 1, at 431.

[9] *Id.* at 433.

[10] *Id.* at 427.

[11] *Id.* at 442.

situation, the uncertainties in identification of the super-norms as well as the chain reactions therefrom are unmanageable.[12]

Obviously, these problems would challenge the three pillars in the international legal system as identified by Weil: voluntarism, neutrality, and positivism. The problem with participation in lawmaking and the expansion of scope of application of law and standing directly challenge the pillars of voluntarism and neutrality. The excessive speed by which this is occurring appears to exhibit a certain amount of overmoralization and thus also challenges neutrality. The uncertainties that may result from all of these problems threaten to destroy the prevailing *lex lata* in the world and, as a result, shake the positivism pillar.

Some of these problems, however, are inherent in the international legal system, even in the classic paradigm of international law. There has never been perfect participation in international lawmaking; the minority always has to wrestle with the majority. The escape hatch, i.e., the persistent objector rule, has also never been a perfect one because, to be able to benefit from it, the relevant state has to be there when the rule originates and then be in a position to make objections consistently going forward. Those who are not privileged to be there first, such as new states or states excluded from the process for whatever reason, or not prescient enough to foresee the new development, may not claim such a benefit. That is to say, perfect voluntarism or complete neutrality has never existed, nor will either ever be possible in international society.

The International Law of Co-Progressiveness as a Response

To these problems associated with relative normativity, Weil's antidotes seem to be a reemphasis on voluntarism, neutrality, and positivism. But the undercurrents for these phenomena were not completely lost to him: "the potential negative consequences of the relativization of international normativity must at worst be regarded as secondary effects of changes that in themselves are beneficial."[13] Still, as a giant who straddled the era of coexistence and cooperation, on the one hand, and the dawning of a new era, on the other, Weil ultimately failed to cross the threshold into that new era.

That new era was waving at me like an impressionistic figure around 2001 when I was attempting to identify the character of international society and international law in the post-Cold War world. I did so by tracing the spirit of society and law at different stages in time, as that spirit may exhibit itself in terms of subjects of the law, content of the law, and enforcement of the law. The *leitmotif* of international law and society was coexistence at the height of the Cold War and cooperation during the period of détente. Since the end of the Cold War, one can see a spirit of society and law that is all-encompassing (in terms of subjects of the law); preoccupied with moral and ethical advancements at an appropriate speed; and internally driven, with human flourishing as its ultimate goal (in terms of the content of the law and the enforcement of the law).

The lawmaking activities now witness the participation of not only traditional subjects such as states and, to a lesser extent, international organizations, but also individuals as well as NGOs. The extraordinary influence of NGOs was on full display in their dramatic success in promoting the conclusion of the Landmine Convention. The all-encompassing character of the participation in lawmaking causes some to abandon the use of the traditional term "subject of law" in favor of "participant in lawmaking" as a general term. In terms of content of the law and its enforcement, what preoccupy us most are no longer the usual issues of coexistence, such as sovereignty, or cooperation, such as economic development, but community interests, human rights matters, and international crime and punishment, thus showing advancements in moral and ethical terms. As there was no preexisting word that follows perfectly from coexistence and cooperation, I had to coin the phrase "co-progressiveness" to describe

[12] *Id.* at 427.

[13] *Id.* at 423.

what I had observed. By the international law of "co-progressiveness" is meant a society or law that is all-encompassing (hence "co"), preoccupied with advancements at an appropriate speed in moral and ethical terms more than in other terms, and having human flourishing as its ultimate goal (hence "progressiveness"). Of course, the *leitmotif* in each period is not the only theme audible to us: in coexistence there was cooperation; coexistence was the background note to cooperation; and coexistence and cooperation are the background notes to co-progressiveness.

I have developed this idea of co-progressiveness in various places[14] and will simply summarize the main tenets here as may be helpful. First, where possible, international law should decide a question with a bent for or a bias in favor of co-progressiveness. Second, progressiveness is to be measured both internally against a given participant's historical achievements and externally across the world against those of other states or participants in the system, but the content and pace of such advancements should be informed by the special circumstances of each state or participant and be ultimately set by that state or participant, limited by the condition that they comply with the most fundamental obligations under international law. This point finds an illustration in the "intended nationally determined contribution" as each state's promised effort under the Paris Agreement to combat climate change. Progressiveness should be usually internally-initiated or self-propelled within each state or participant; it can also be externally induced—but not coerced. This point would not endorse hard conditionality in economic assistance programs that would impose a certain course of governance reform, but it would counsel in favor of soft conditions such as requirements that recipients of assistance participate in educational programs that would expose them to best practices while leaving to them to decide themselves whether or not to adopt them.

Third, clashes between intrinsic and instrumental values should be decided by decision-makers in a conscious and explicit way and by giving preference to a more important value but, at the same time, to the applicable intrinsic value where possible.[15] Fourth, every state or appropriate participant (including international organizations, regional organizations, and perhaps individuals) in the international system is a holder of rights and bearer of obligations vis-à-vis each other as well as vis-à-vis the international system or the international community of common interests or community of shared future for mankind. Fifth, the equality to be pursued should be enlightened equality, not mechanical or superficial equality. It should be based on a fitting and progressive criterion in each instance, appropriate to the particular subject matter at issue and the occasion, so as to make meaningful or sophisticated the twin maxims that "like cases should be treated alike" and "different cases should be treated differently." Sixth, great states and leader states enjoy special powers and thus should shoulder special responsibilities in the international system. Seventh, every state or appropriate participant in the international system is to observe the rule of law in every respect and, in particular, to take the best account of rule-of-law concerns in making every important decision, especially the need for taking a rigorous approach thereto.

Born out of my (perhaps rosy) observations of the spirit of international society and law, the idea of the international law of co-progressiveness was not intended as a direct response to the problems that Weil identified. Nevertheless, it does come with a toolkit full of responses that would go a long way to solving those problems or at least reducing them to a minimum.

As to those problems inherent in the international legal system, such as imperfect participation in lawmaking and the perennial struggle between the majority and the minority, the international law of co-progressiveness cannot eliminate them, either. Its promotion of all-encompassingness and greater participation in the legal system (including lawmaking), however, would no doubt reduce the feeling among many states of being left out. Still, the need to maintain an international community at least with respect to the most important matters, i.e., *jus cogens*,

[14] *See supra* note 2.

[15] *See* YEE, PART II, *supra* note 2, at 85.

requires the elimination of the escape hatch, i.e., the persistent objector rule, in such matters. The alternative to this rule would be to prioritize one particular state over the entire community, a scenario that is not appetizing, either.

In promoting a perspective of shared progressiveness in all decision-making, the international law of co-progressiveness would ultimately bridge the gap between different ideological orientations and thus alleviate or even eliminate the challenge to neutrality. Furthermore, it does so by relying on internally driven progressiveness within each state or participant in the system, even if induced (but not coerced) by outsiders; this would promote voluntarism on the part of states and participants and reduce the feeling of having views imposed on them.

The backlash against problems with excessive speed and overmoralization has been driven home to us all by the rise of particularism, local boosterism, antiglobalization, or a retreat from international institutions. The international law of co-progressiveness recognizes that one cannot rush things too much and thus emphasizes that progressiveness must be achieved at each participant's own appropriate pace, not to the satisfaction of all but better than no progress at all.

This law also considers that the addressees of the need to be co-progressive include all states and other participants, strong or weak. In this regard, this law recognizes the enormous impact that strong or leader states[16] may exert, whether in promoting progressiveness by providing constructive engagement with other states and participants, or in wreaking havoc in the world by using their advanced positions as a tool of oppression. When the latter happens, one will see overmoralization or self-interest taking reign to the extreme. Obviously, these states or participants themselves need to undergo a heavy dose of co-progressiveness, so as to avoid severe backlash or a "tooth and nail" fight back from the oppressed. Here the intelligentsia in each state or participant would play an indispensable role.

To all the problems, especially the problem with uncertainty, the most potent antidote is the strict application of the rule of law, particularly in decision-making, that the international law of co-progressiveness emphasizes. The championing of community interests and the increasing recognition of the standing of states in the enforcement of community interests (to the extent they are so recognized by the international community) are to be celebrated. The tough task is to identify and apply the exact extent of such championing and recognition. In this regard, the international law of co-progressiveness demands evenhandedness, condemns double standards, and imposes a rigorous approach to decision-making on all relevant decision-makers, especially legal decision-makers such as judges and arbitrators. Too often we see decision-makers overrecognize such community interests and do so by taking a fast and loose approach to decision-making, such as skipping many pivotal decision-points in order to achieve their goal. Such decision-making does a disservice to, rather than helps the cause of, community interests. These "overachievers" give lawyers a bad name.

If a rigorous approach to decision-making does identify the extent of the recognition of a community interest and its associated regime, such as *actio popularis*, its application probably would not create the kind of chaos that Weil feared. For example, the International Court of Justice accepted a kind of *actio popularis* or obligations *erga omnes partes* within the context of the Torture Convention,[17] and is confronted with the same issue in the context of the Genocide Convention in an ongoing case.[18] *Actio popularis* pure and simple may one day appear. One may ask whether, on the issue of standing, the Court has conducted a rigorous exercise so far. If *actio popularis* in the context

[16] On the role of strong states and leader states, see *id.* at 123.

[17] *See* Questions Relating to the Obligation to Prosecute or Extradite (Belg. v. Sen.), 2012 ICJ REP. 422, paras. 68–69 (July 20). The standing issue potentially was present in Whaling in the Antarctic (Austl. v. Japan: N.Z.), 2014 ICJ REP. 226, but it was not discussed by the Court.

[18] *See* Application of the Convention on the Prevention and Punishment of the Crime of Genocide (Gam. v. Myan.), I.C.J., https://www.icj-cij.org/en/case/178.

of a treaty or even pure and simple is the result that a rigorous approach to decision-making would in fact give us, a court or tribunal's exercise of this approach may help the respondent accept it and convince members of the international community that being taken to court is just a normal part of the burdens of being a member of that community. The respondent may take measures to limit consent to jurisdiction by excluding *actio popularis* cases. Or perhaps measures can be taken to compensate a respondent prevailing on the merits, so as to prevent the abuse of the *actio popularis* regime. Regardless, a rigorous approach to decision-making would help to promote the steady progress in recognizing community interests.

Conclusion

This short discussion gives one the feeling that the various problems associated with relative normativity are enduring problems. As a result, the efforts to counter them will probably have to be enduring efforts, if not Sisyphean. Still, the international law of co-progressiveness as highlighted here and seen as a toolkit points in a good direction. The world of coexistence and cooperation, plus co-progressiveness, will be a better one, despite all the rough-going in the world today.

BMC Palliative Care

RESEARCH

Open Access

Feeling called to care: a qualitative interview study on normativity in family caregivers' experiences in Dutch home settings in a palliative care context

Maaike M. Haan*, Gert Olthuis and Jelle L. P. van Gurp

Abstract

Background: Family caregivers, such as partners or other family members, are highly important to people who desire to stay at home in the last phase of their life-limiting disease. Despite the much-investigated challenges of family caregiving for a patient from one's direct social network, lots of caregivers persevere. To better understand why, we aimed to specify how normative elements – i.e. what is considered good or valuable – shape family caregivers' experiences in Dutch home settings.

Methods: From September 2017 to February 2019, a total of 15 family caregivers, 13 bereaved family caregivers, and 9 patients participated in one-time in-depth interviews. The data were qualitatively analyzed following a grounded theory approach.

Results: Central to this study is the persistent feeling of being called to care. By whom, why, and to what? Family caregivers feel called by the patient, professionals entering normal life, family and friends, or by oneself; because of normative elements of love, duty, or family dynamics; to be constantly available, attentive to the patient while ignoring their own needs, and assertive in managing the caring situation. The prospect of death within the palliative care context intensifies these mechanisms with a sense of urgency.

Conclusions: Our analysis showed a difference between feeling called upon in the caring situation on the one hand, and how caregivers tend to respond to these calls on the other. Taking into account the inherent normative and complex nature of family caregiving, the pressing feeling of being called cannot – and perhaps should not – simply be resolved. Caring might be something families just find themselves in due to being related. Rather than in feeling called upon per se, the burden of care might lie in the seeming limitlessness to which people feel called, reinforced by (implicit) social expectations. Support, we argue, should enable caregivers to reflect on what norms and values guide their responses while acknowledging that caring, despite being burdensome, can be a highly important and rewarding part of the relationship between partners or family members.

Keywords: Caregivers, Family care, Informal care, Palliative care, Ethics, Caregiver burden, Qualitative research, The Netherlands

*Correspondence: Maaike.Haan@radboudumc.nl
Radboud University Medical Center, Radboud Institute for Health Sciences, IQ healthcare, P.O. Box 9101, 160, 6500 HB Nijmegen, The Netherlands

Background

In the last phase of a life-limiting disease, patients are often cared for by someone close from their social network, such as their partner, grown-up child, or friend.[1] These people providing family care (Table 1) are pivotal. Due to their often unique relationship with and valuable knowledge about the patient, family caregivers are essential in providing emotional support, communicating with professionals and services, relieving pain and other symptoms, or doing practical tasks [9]. However, especially if the patient desires to stay and die at home, as most people prefer initially [10], the roles and responsibilities of their families and friends are intensified [7, 11] and caregiving may be burdening. This study explores why family caregivers (hereafter: caregivers) persevere, despite the challenges, and which role normativity plays.

Previous research has highlighted the physical and psychosocial challenges of family care in the palliative phase, recognizing the need for caregiver support [12–14]. Many caregivers live in permanent uncertainty about the future and feel overwhelmed and unprepared for their caring role [15, 16]. Maintaining normality in social engagements can be a struggle, for instance when a caregiver's sense of togetherness with the patient conflicts with also feeling socially isolated [7, 17]. Caregiving can limit or even 'chain' caregivers in their own life and affect their relationship with the patient [16]. About one in five caregivers of terminally ill patients experience a heavy care-related burden [8]. This much-investigated concept of 'burden' is reported to be associated with factors like distress at witnessing suffering and disease progression, uncertainty about the situation, role strain, sleep deprivation, spiritual distress, and financial crises [1, 7].

As lots of people feel burdened by caring for someone close, why do they persevere? Many empirical studies focus on predefined outcomes, but little tell us about the individual processes and context that shape the actual care [18]. Care ethical analyses, rooted in feminist studies, have argued that caring is not (only) a matter of one's free and personal choice. Rather, a caregiver's agency is deeply tied to the surrounding social and political practices: *"caregivers appear as people finding themselves in a position in which others, they themselves, and also the socio-political context* expect them *to have and take responsibility, as a result of socio-political, personal, affective, contextual, and ethical factors."* (p. 277) [19] In this article, we are specifically interested in these ethical factors, which we label as 'normative', e.g. the normative elements that appear to be essential to the experiences of family members providing care. Normative elements have to do with someone's convictions about what is 'good' or 'right' to do, which are particularly relevant when it comes to life and death in the last phase of a life-limiting disease. As Randall and Downie (p.13) argue:*"... any discussion about palliative care occurs against the background of those major questions which relate to the meaning of life and death, or what constitutes a good life (and perhaps death) for a person."* [20] Normativity has already been recognized as important in palliative care and in family caregiving, but its precise role remains unclear in the caregiving literature. For example, the theoretical Informal Care model proposed by Broese van Groenou and De Boer (2016) understands family caregiving, in general, as being dependent on both contextual factors that influence care provision as well as individual factors that shape one's intention to provide care, to which they refer broadly in normative terms, i.e. beliefs, motives, values, or norms. Prior empirical studies in palliative care also refer to normative caregiver experiences, such as feeling unprepared yet *responsible* to provide care [15], feeling a general *duty* to care [21], or feeling *obliged* to prefer providing care at home [22].

Drawing on these broad references to normativity, we aimed to specify how normativity shapes the phenomenon of family caregiving within the context of palliative home care. This article's research question is: how do family caregivers of seriously ill people in Dutch home settings experience caring for their partner or family member in the palliative phase, and how are these experiences shaped by normativity? More insight will provide us with relevant clues about how to better understand and support caregivers.

Methods

Study design

Our overall research project aimed at exploring the diverse palette of family caregiver experiences, following the characteristics that are key in the diverging views and philosophical assumptions of a grounded theory methodology [23, 24], i.e. we used an inductive approach, we simultaneously collected and analyzed data while developing theoretical abstractions grounded in the data, we used constant comparison and kept memos, and strived for theoretical sampling. Thus, we ensured a thorough exploration of people's experiences with caregiving and remained open to both the positive, rewarding experiences that come from caring as well as the experiences of feeling burdened [8]. In line with Straussian grounded

[1] We as authors are aware of the debates about the use of the term 'patient'. We acknowledge that, especially within the context of this study, people are foremost each other's partner, parent, friend, or other family member. For the sake of clarity, however, we use the more formal term patient for the person who receives family care.

Table 1 Definition and background of family care in the Netherlands

In this study, family caregiving is regarded as the wide range of aid or assistance in activities of daily living given by an unpaid and untrained person from someone's direct social network, e.g. a partner, relative, grown-up child, friend, neighbor, or other acquaintance [1–5]. Family care may vary in intensity and duration, but, in any case, goes beyond what can reasonably be expected within the relationship [4]. In the Netherlands, several definitions and criteria for 'family care' exist, leading to different estimations of the number of caregivers [6]. Defined broadly, about 35% of people of 16 years and older reported providing unpaid help to someone close with health-related problems [6]. Although the terms are sometimes used interchangeably, we deliberately do not use the term 'informal care' to explicitly exclude volunteers from this study.
Family care occurs in several settings, i.e. for people suffering from serious illness, long-term mental health problems, or a disability. This study specifically focuses on family care in a palliative care context, as a consequence of metastasized cancer or long-term organ failure. Estimations concerning the number of caregivers in this context are difficult, especially if hand-on help from members of the wider social network, other than the 'primary' caring relative, is taken into account [7], which occurs in the majority of Dutch caregivers of terminally ill patients [8]. The lack of a clear-cut point of entering the end-of-life phase further problematizes the estimation [7, 9], as does the observation that not every 'family caregiver' recognizes himself or herself as such [3].

theory [23], pre-given concepts or previous studies did not dominate the data collection and analysis to keep an open mind [24–26]. In line with Charmaz' constructivist approach, however, we considered ourselves as not having an empty head [25]. We started the project with some sensitizing concepts (that were adapted into topics for our interview guides, see Table 2), based on an explorative search of scientific and gray literature, as well as interview expertise in palliative care within the project group (JvG, GO).

During the cyclical data collection and analysis, our background as medical ethicists made us notice normative aspects in the interviewees' phrasing and overall stories. We became increasingly interested in what motivates caregivers to care for patients in need of palliative care. Therefore, for this article, we specified the initially broad research question of the overall project into how the collected care experiences are shaped by normativity.

Setting, participants, and materials

Based on the inclusion criteria (Table 3), three groups of Dutch interviewees were included: 1) family caregivers, 2) bereaved family caregivers (within 5 months after the patient's death) who could describe the last days before and first few weeks after the death of the patient, and 3) patients i.e. people with a life-limiting disease in the palliative phase who received family care. All interviewees

Table 2 Topics in the interview guides

Caregiver interviews:
- Situation and disease process of the patient, (changed) relationship with the patient
- A (non) typical day and caregiving
  - In case of bereaved caregiver: last weeks and time after the patient's death
- Meaning of relationships with others and support
- Taking care of oneself, needs, support
- Saying goodbye, talking about death, future

Patient interviews:
- Situation and disease process, meaning of being ill, (changed) relationship with the caregiver
- Being cared for, a (non) typical day
- Meaning of relationships with others
- Saying goodbye, talking about death, future

Table 3 Inclusion criteria

All interviewees:
- are 16 years old or older;
- are mentally competent to engage in the interview;
- are fluent enough in the Dutch language to participate in the in-depth interview;
- are involved in family care, as either caregiver or care receiver;
- are aware that the involved patient is in the palliative phase of life and has a limited life expectancy.

The involved patients (not necessarily interviewees):
- are 16 years old or older;
- have an incurable and life-limiting oncological or neurological disease, organ failure, or elderly frailty (dementia is excluded, because of the specifically changing nature of the relationship between caregiver and patient);
- are in such condition that their involved healthcare professional would not be surprised if this patient died within the next 12 months (e.g. the *"surprise question"* as used in palliative care);
- receive family care at home from at least one loved one (partner, child, parent, sibling, friend, etc.); or has received this care before being transferred to a hospice or other care institution.

were aware that the trajectory of the involved patient was regarded as palliative. We were focused on caregivers' perspectives primarily. Patient interviews were considered to be complementary data, functioning as triangulation, to better interpret and understand the context of the caregivers' data. Purposeful sampling was used to include a variety of people, based on gender, age, or caring relationship [24]. Various healthcare professionals (e.g. general practitioners, district nurses, hospital-based professionals) invited potential interviewees, leading to convenience sampling as well. Later on in the study, we strived for theoretical sampling to identify variations and relations in the ongoing analysis [24]. When a family caregiver or patient was interested, MH contacted him or her via telephone, and sent an information letter. The interviews were scheduled within days or weeks after MH first contacted the interviewees.

The interview topics (Table 2) and questions were reviewed by various experts in qualitative research and palliative care. We constructed different guides for the different interview groups, that were adapted several times during the data collection to deepen the ongoing analysis and achieve saturation. First author MH received interview training from an independent senior researcher with ample experience in qualitative research and personal experience with family caregiving. The interviewing was piloted with the trainer twice.

Data collection

MH conducted the interviews. Although the interviews were guided by general topics (Table 2), they were minimally directive. Probing was mainly based on the situation and what the interviewee said. Field notes and memos were made to facilitate the iterative process of interviewing and analyzing. The interviews were audio-recorded and transcribed verbatim. The original Dutch quotes for this article were translated into English.

Ethical considerations

Ethical approval was sought from the Medical Review Ethics Committee region Arnhem-Nijmegen (registration number 2017–3415), who determined that this study does not fall under the scope of the Medical Research Involving Human Subjects Act (WMO). Potential interviewees verbally consented to be contacted by MH. All participants gave their written informed consent, after receiving an information letter. Because healthcare professionals acted as our gatekeepers in finding participants, the potential interviewees' privacy was an important point of consideration. In addition, given the interviewees' vulnerable position, caution was needed in inviting and interviewing people. The possible emotional or physical burden of the interview was acknowledged both in the information letter and the interview itself; people could withdraw at any time. Furthermore, because the interviews involved personal and emotional topics and we did not want to disturb existing relationships, MH tried to conduct the interview individually without others nearby – unless the interviewee wished differently or practical circumstances did not allow this. Every participant received a numerical code and all interview transcripts were anonymized. The data were safely stored, as was stated in an approved data management plan.

Data analysis

Following an iterative approach, characterizing grounded theory, the data were analyzed throughout the collection process [27, 28], using ATLAS.ti software. MH first coded transcriptions individually and constantly compared the analysis to previous codes, staying close to the language used by the interviewees [25]. As the cyclical process of grounded theory allows, rereading, coding, and collecting new data occurred simultaneously to dig deeper into the research question [25]. In line with Straussian grounded theory, open, axial, and selective coding were used to find categories, overarching themes, and, ultimately, patterns [23, 24]. Meanwhile, hunches and decisions concerning codes, categories, and patterns were kept in memos [26, 29] to add to the transparency and credibility of the research. Several visual displays were made to organize ideas and discuss the theory [29]. The data collection continued until new interviews no longer provided new insights on general patterns within the group of caregiver participants. Although the use of theoretical sampling was limited, we then decided that saturation was reached.

Analytic rigor

In contrast with a classical view on grounded theory [23, 29], we followed Charmaz' constructivist approach in regarding the researcher as not being neutral or distinct from the research process [25] and stressing the importance of a researcher's reflexivity conducting grounded theory [23]. Thus, for validity reasons, JvG independently coded the first three interviews as well [25, 27]. We critically discussed our differences as coders without looking for a compromise. During the interviewing, MH verified her interpretations with the interviewee by summarizing and choosing multiple angles during the interviews. General patterns and themes were checked to employ specific interview questions in later stages of the data collection. Moreover, in June 2018, the identified themes and patterns in the data analysis were discussed during a feedback meeting with five interviewees. In addition, the authors regularly discussed their results with a sounding board of experts, mainly from various stakeholder

organizations. Utilizing memo writing, but also discussion and checks both within and outside the project team, measures were taken to ensure MH's reflexivity and critically consider her role and possible biases (for example, focusing on challenging aspects of caregiving in the interviews). Throughout the study, ongoing peer review and discussion between all authors were essential.

Results

From September 2017 to February 2019, a total of 15 family caregivers, 13 bereaved family caregivers, and 9 patients participated in one-time in-depth interviews (Table 4). All caregiver interviewees were (one of) the involved patient's primary family caregivers. Six interviews were conducted with two interviewees present (caregiver and patient, or two caregivers), due to their preference or initiative. In all interviews, the main subject of conversation was family caregiving for a patient in the prospect of an approaching death. Life expectancies of the involved patients varied from weeks, months, to maximally a few years. Most interviewees were interviewed individually at home or where they resided; two caregivers preferred to be interviewed without the patient nearby and were interviewed at the researchers' department. Interview duration ranged from approximately 38 min to 2 h and 27 min.

Overview

Our results present a qualitative analysis of the phenomenon of family caregiving in palliative care, as seen from the perspectives of various caregivers and patients, focusing on how normativity shapes caregiving behavior (Fig. 1). Central

Table 4 Participant characteristics

Characteristics	**Family caregivers (*N* = 28)**	**Patients (*N* = 9)**
Male / Female	12 M / 16F	4 M / 5F
Age	23 to 84 years (mean 58.1)	61 to 95 years (mean 74.2)
Retrospective interview	13	not applicable
Relationship between patient and primary family caregiver		
• Partner	18 partners	7 partners[a]
• Child–parent	9 children (1 son)	1 parent[a]
• Other family member or friend	1	2
Diagnosis		
• Cancer	16	5
• Organ failure (e.g. chronic obstructive pulmonary disease or heart failure)	6	2
• Other, comorbidity or unclear condition	6	2

[a] One of the 7 partner interviewees also received care from her child and specifically told about that in the interview

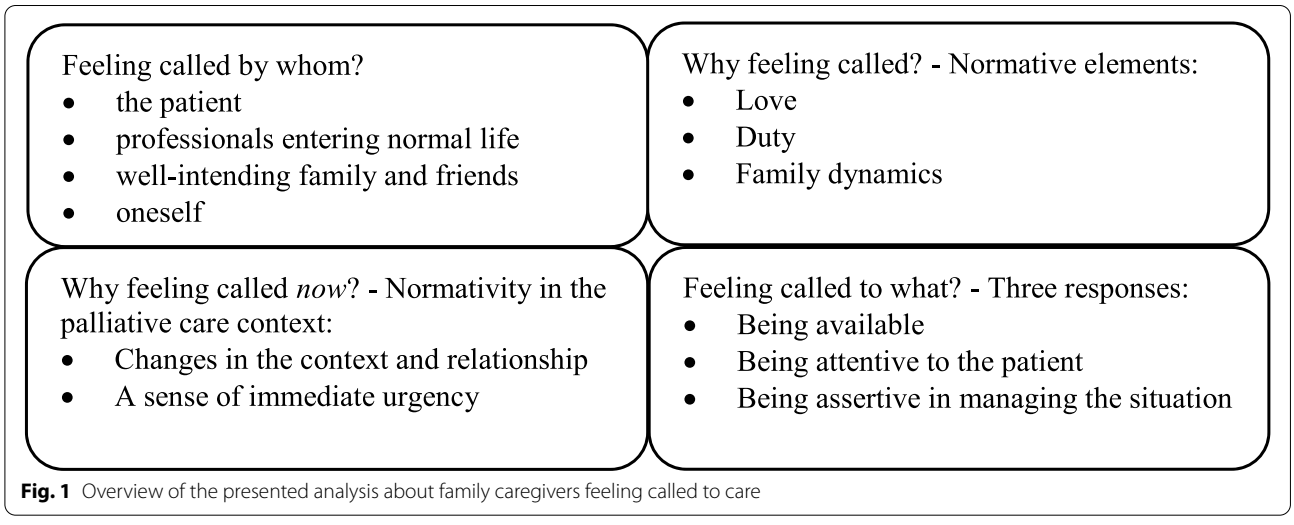


Fig. 1 Overview of the presented analysis about family caregivers feeling called to care

to this study is the persistent and sometimes continuous feeling of *being called to care*. Our analysis further explains a pattern in this feeling: by whom, why, and to what?

Caregivers feel called either by the patient's explicit or implicit calls for help, professionals entering normal life, well-intending family and friends, or by oneself. Their responses to feeling called seem to be evoked by general normative elements, e.g. love; duty; and family dynamics. More specifically, by subjecting relationships to pressure and intensifying the felt calls with a sense of urgency, the palliative care context seems pivotal in understanding to what caregivers feel called, e.g. being constantly available, attentive to the patient while ignoring their own needs, and assertive on several fronts in managing the caring situation. Our analysis, thus, revealed a difference between feeling called upon in the caring situation on the one hand, and how caregivers tend to respond to these calls on the other – reinforced by normativity.

Feeling called by whom?

We found a persistent and always-existing urge which we identified as the manifestation of *feeling called to care*, this study's central theme. Looking at the caregivers' stories from an ethical perspective, this feeling presents itself as pre-reflexive in their stories, e.g. not based on a well-considered choice using explicit moral arguments: caregivers often felt a strong urge to act and automatically did so, yet without having been able to thoroughly reflect on why or how. Some people reported feeling what was needed intuitively because they knew the patient so well and the two of them spent so much time together. Feeling called to care manifested itself in the caregivers' experiences in four ways: by the patient; professionals entering normal life; well-intending family and friends; or one-self.

Feeling called by the patient The family caregiver was often the person closest to the patient. Caregivers felt called by patients explicitly, when they needed something to eat or drink, asked for assistance with going to the toilet, expressed anxiety and worries, felt ill, or even screamed in pain. Patients were fearful and stressed when they were left alone, some patient interviewees tried to maintain in control by having (multiple) phones within reach, frequently asking or calling their caregivers. As a result, caregivers experienced a, sometimes continuous call:

> *... I have the feeling that I am being deeply called upon to do so, that it is 24-hour care. Really. [...] It starts in the morning when I get up. The first thing I do is put her on the bedpan. [...] It goes on until she goes to bed; it is continuous. Sometimes you can have 15 minutes or so, a few minutes, or if someone is there, you don't even have that. But it is actually always there. [...] Every moment is an appeal, every moment. ... (i17, husband caring for his wife with a lung disease)*

This was endorsed in retrospective interviews: some bereaved caregivers felt relieved by not having to continuously respond to the urgent needs and suffering anymore, which showed the urge of feeling called explicitly.

Next to explicit calls, patients showed implicit calls, for example in their preference for and dependence on the family caregiver being the primary person to provide care:

> *Patient: Family care is everything here right now. Without it, there is nothing left. [...] Interviewer: Why do you think it's so important that she [his wife] gets more help? Patient: So she can carry on, eh? For herself. For her daughter. It sounds selfish, but also for me, eh? Because if she fails ... Tell me ... (i8, male patient with cancer being cared for by his wife)*

Feeling called by professionals entering normal life In caring at home, caregivers also had to deal with professionals, technology, or aids. Usually, they were untrained medical laypeople, feeling overwhelmed by disease symptoms:

> *... he was in pain and felt he couldn't get air. I have no medical background or anything, so. [...] It is like being in an unfamiliar forest, so you don't even know what is there. (i11, wife caring for her husband with cancer)*

Professional home care, then, could offer relief by taking over caring tasks and responsibilities. At the same time, professionals' presence did not seem to resolve the aforementioned patient's calls. People would still feel alert and called to provide care themselves:

> *Daughter: And the best thing about it [interviewer name], that my father called me after all. "[Daughter's name], where are you?" So I had to get up. Because I heard my father, and then the girls [from the night care] said, "no, just stay put". I say "yes, my father calls me, never mind. I'm awake anyway", I say.*
> *Son-in-law: Yes, and he does not understand Dutch well either.*
> *Daughter: No, and then he started talking [foreign language] to those girls too. (i21, bereaved daughter and her partner caring for her father with cancer)*

In addition, the presence of professionals instigated a call itself in several ways. For some people, it was difficult having strangers in their home several times a day, taking-over normal life: the *"circus"* of professional caregivers stopping by according to one interviewee. Caregivers entered what an interviewed son said to be the *"new world which is called care"*, sometimes accompanied by bureaucracy. Negative experiences with institutional care or the tight schedule of professional home caregivers also urged family members to provide the care themselves. Being able to organize care in the way one chooses (e.g. time or place) was seen as an advantage of providing family care instead of waiting for professional help.

Feeling called by well-intending family and friends Caregivers and patients were often surrounded by family or friends, by whose attention they felt supported and loved. Despite being well-intended, their visits, how-are-you-questions, worries, or advice could instigate a call for caregivers, notably when mediated by social communication technologies demanding a continuous presence. Some patients found it difficult to maintain their autonomy. Having to tell the sometimes confronting truth over and over again could be difficult for caregivers. Yet, according to some, by explaining the situation regularly, it also lost some of its emotional edges. In a way, the disease seemed to take over normal life and conversations:

> *But often you also have to talk about the disease before you can talk about other things. So when I join in – I play tennis once a week with a friend, a tennis mate. First I talk to him for 10 min about [wife's name] and the disease. Then we can play tennis. [...] And if you have had that, say, once you have bitten through the sour apple, then you can also enjoy the rest. (i14, husband caring for his wife with cancer)*

In some family contexts, talking together about death was taboo, resulting in a tension between respecting one's family norms vs. one's own desire to discuss last wishes, the upcoming funeral, or other end-of-life issues.

Feeling called by oneself Lastly, feeling called to care could be self-imposed. Caregivers sometimes felt like they were the only, or the most capable, person in the family to arrange things, wanted to do things themselves, or specifically asked what the patient wished:

> *Yeah, maybe I'm in that [appeal] too, because I'm going to ask him, "what would you really like to eat?" [...] That appeal, I pull it towards me. It is not that he says "you have to do that for me". I ask him what he would like. (i29, son caring for his father with aneurysm)*

Why feeling called? – Normative elements

What motivates people to respond to these calls? This section explores normativity essential to the caring situation and relationship with the patient. We found several, often co-existing, normative elements that evoked caregiver responses: love, duty, and family dynamics.

Love First, caregivers felt called to care for the patient out of love or a special bond:

> *Yes, you just do that for each other. [emotionally] He is my husband you know, so I love him. You don't want someone else to do all that for you. (i28, wife caring for her husband with cancer)*

Caring out of love could be based on a promise, for example between partners (for better or worse) or other family members. Caring was also seen as self-evident in a loving relationship or was valued because it enabled people to live their normal life as much as possible. One patient noticed that being cared for by her partner was different from being cared for by her son because the latter was less evident (*"he is still my child"*). According to several daughters, love is what makes family care different from professional care. Caring could also be expected because the patient only trusted family members, or it stemmed from reciprocity in a loving relationship: for some, it was important to imagine themselves in the situation of the patient to make decisions. However, siblings sometimes had conflicting opinions about what was best or most pleasant for the involved parent.

Duty Caring out of love sometimes co-existed with feeling a duty, for example keeping a promise regardless of whether love was still felt. Care could also be provided out of obligation or duty in the relationship or in the wider family:

> *..., No, I don't have to be put in the limelight because this is just a part of my job. It's a piece of moral duty that belongs to the fact that I am a daughter of my mother who took care of me. (i13, bereaved daughter caring for her mother with comorbidities)*

This obligation stemmed from social expectations concerning familial relationships or helping others in need.

Important to note here is that we only interviewed caregivers in the Netherlands and did not specifically ask for demographic information about culture or religion. Amongst interviewees who described their culture as family-oriented, we found a moral duty to take care of one's parents in return for their upbringing. An interviewee, however, emphasized love and honor as motivating in caring for her father, instead of this norm of reciprocity or gratitude in her Islamic culture. The value of helping others in need, which was sometimes religiously inspired, was also found amongst interviewees with an intuitive intention or moral obligation to help whenever they could.

Family dynamics Especially in parent-child relationships, the sometimes troubled family history and the specific relationship with the parent played a role in what caregivers expected of themselves and how they responded in the caring situation. For example, being the most responsible sibling, being the darling sibling of a mother, or the past conflicts between family members. In addition, contrast experiences with another family member's death or with caregiving in the past were directive in persevering:

> *… "I will take care of him until I drop, and then I can't do it anymore", I say. This will never happen to me the way it did to my mother. [...] Fortunately, I have managed to accomplish it. (i19, bereaved wife caring for her husband with a neurological disease)*

A family's hierarchy also influenced caregivers' feeling of being called to care. Having a specific position within a family, such as being the eldest or being a professional healthcare provider, could lead to being assigned to the role of primary caregiver by one's siblings.

Why feeling called now? – Normativity in the palliative care context

To some extent, the aforementioned calls and normative aspects apply to all types of long-term family caregiving at home. What, however, is markedly different in a palliative care context is the décor of deterioration and the prospect of an approaching death. This section explores normativity within this specific context.

Changes in the context and relationship Our results showed the context to be changing and increasingly palliative, consisting of a poor prognosis and deterioration, i.e. physical changes and/or a loss of autonomy while handing over things and becoming increasingly dependent on care. Some patients feared dying alone, choking or in serious pain, or worried about what they had to leave behind. In partner relationships, the inevitability of death could lead to feelings of loneliness, being unable to live and share (sexual) life as they used to, or having to find proximity in other ways, and no longer sharing a joint future. The deterioration and fear seriously-ill people experienced increasingly changed them into a 'patient' in the eyes of their partner or family member: the person slowly *"faded"*, for example, and was no longer always recognizable as he or she used to be. These changes invoked a role shift for the partner or family member as well, who became more or less, sometimes in an unwanted way, a 'caregiver' with new responsibilities:

> *Yes, no, you don't want that. You just want your partner to be your partner. Often that's what she is too. If I have things to deal with, then she is still the person I go to, the one who gives me advice. I always want to hear what she thinks about it. We can level with each other, we can spar with each other. But there are just considerable times when I am home help for her. (i14, husband caring for his wife with cancer)*

A sense of immediate urgency The prospect of nearing death, which changes the relationship, also leads to a sense of urgency in responding to felt calls: caregivers believed they had one final chance to do the right thing for the patient. This seems pivotal in understanding what motivates them to immediately respond to calls, sometimes regardless of whether this is rewarding or physically or emotionally burdening:

> *…that's what I said, I only have one chance of doing this for him. And that is simply why I am putting my back into it and why I am doing it. (i29, son caring for his father)*

Feeling called to what? – Three responses

Our results showed that the aforementioned normativity within the context urges caregivers to be constantly available; attentive to the patient first while ignoring their own needs; and assertive on several fronts of their lives in managing the caring situation. Despite the dilemmas described below, it is important to note that intensively caring for the patient also evoked *positive* feelings among caregivers. Although bereaved caregivers acknowledged exhaustion, they spoke of being "grateful" for having persevered and having enjoyed precious moments together, feeling *"honored"* or *"proud"* to have been able to provide family care. Some spoke of a closer and more intense relationship, for instance by spending more time together

and cherishing small things or meaningful moments as the *"last things"* in the light of the approaching death.

Being available In reaction to both the urgency and insecurities in a worrisome palliative care situation, as well as the patient's fear of being left alone, caregivers strived to be always available, reachable, and alert – live or by phone. Some caregivers felt *"imprisoned"* in their home, feeling obliged to be constantly stand-by for help. Leaving the house was easier when someone else stayed with the patient. Yet, overall, caregivers found it hard to let go of the situation and not think or worry about it, even when the patient was out of sight or when they themselves were out of the caring situation for a moment to get groceries, see friends or family, take part in sports, or work, especially with other people's questions. This illustrates the pressure of feeling called: despite leaving the situation, caregivers still felt occupied. Some bereaved caregivers experienced feelings of guilt for not having been present all the time:

> *Interviewer: … You said that it was painful for you that you were not there with him at the last moment. Sister-in-law: Yes, because I left him alone […] while I knew he was so afraid. Because I assured him, "[name of brother-in-law], when the time comes, I will take care that I am there with you." But this was unforeseeable. (i22, bereaved sister-in-law caring for a male patient with chronic obstructive pulmonary disease)*

Being medical laypeople, caregivers were sometimes unable to take away pain or breathlessness. Some told of situations in which they stood by helplessly or actually preferred to not see the patient's suffering, which also shows the tragedy of being present all the time.

Being attentive to the patient first We found a persistent tendency among caregivers that everything should be about the seriously ill patient, leading to self-ignorance, e.g. meeting the needs and wishes of the patients first, before addressing their own. The disease literally took over caregivers' normal life in their own house: medical aids on the dinner table, a bed in the living room, or shelves full of medication. Work, volunteer jobs, social activities, or trips were put on hold in light of care or spending time together. Caregivers sometimes felt lonely in their daily struggles, or frustrated about the lack of attention from professionals for their perspective:

> *… but a family doctor like that who never says, "How are you doing? Can you manage?" […] you think it's normal, you don't know any better. You think, it's not about me; it's about papa. All the while I was the pivot. (i31, bereaved daughter caring for her father with cancer)*

Caring could result in physical exhaustion, but several people also pointed to what we have identified as a normative pressure stemming from being called in light of an inevitable death:

> *Yes, it was tough, really tough. And I don't mean that I was physically tired, but you reach a point where you are in a sort of tunnel, and you just go on and on. Well, also later on, you think 'my God, phew.' But there is simply no going back from there, in the sense that you want closure in a proper way. (i25, bereaved husband caring for his wife with cancer)*

The feeling that one can never turn back the clock again was said to be motivating:

> *… I can get up in the morning […], I look in the mirror, I feel no guilt because I have done everything, and you can't turn back the clock. So, in a loved one's last phase, you have to try to do everything you can, try to do what you would like to do. Because after that, you can't do anything, and you could regret that. (i21, bereaved daughter caring for her father with cancer)*

Balancing attentiveness to the patient with leading one's own life could cause dilemmas for both patients and caregivers. Patients sometimes worried about being a burden for their families, acknowledging their needs and preferring to not restrain them in their activities outside the caring situation. Some caregivers would not express their burden to the patient to spare him or her from feeling guilty. Having to raise young children but not being able to share this anymore with a seriously ill partner, could further problematize finding a balance. Overall, patients, as well as caregivers, seemed to want to respect each other's wishes, even if this conflicted with their own, for example about (not) having open conversations about death, arrangements, or last wishes.

Societal expectations played a role in some dilemmas. On the one hand, going out and engaging in fun activities could lead to a feeling of having to explain one's behavior to other people, showing the persistent norm that everything should be about the patient. Bereaved caregivers sometimes experienced feelings of doubt, seeking confirmation that they had done it right and had given enough. On the other hand, a caregiver's large amount of time invested in caring or a self-sacrificing attitude could evoke critical questions by friends or family as well:

> *To fight the resistance of others [raises her voice]:*

> *"Oh, it's irresponsible." "You are sacrificing yourself and you have nothing left." […] It felt as if I had to justify everything while I know very well what my responsibility is. My responsibility is not somebody else's responsibility, so there. I know what's best for him, what's the most fun, what he likes most. ... I focus it all on that. (i19, bereaved wife caring for her husband with a neurological disease)*

Being assertive in managing the situation While being constantly available and attentive, caregivers felt demanded to uptake an assertive attitude on several fronts of their shared lives, as a response to the diverse calls in the caring situation.

Concerning practical caring duties, caregivers often referred to themselves as *"manager"*, e.g. doing household tasks, preparing and sometimes administering medication, accompanying the patient to doctors' appointments, informing family and friends and managing visits, maintaining contact with organizations and healthcare professionals, arranging for medical aids and adjustments to the home setting, or doing the record-keeping and managing accounts. They also had to assertively call for help in time. Sometimes, a *"battle"* was fought with organizations or professionals, with caregivers being the patient's spokesperson.

Assertiveness could be required in the caring relationship itself too, in confrontation with deterioration. Some partners motivated their loved ones when they themselves no longer could, made jokes or had a laugh to make it bearable:

> *Because I was relatively stronger than the rest of the nurses. I picked her [his wife] up [to put her on the bed pan] and she went with me. I said "Let's go dancing", and then I maneuvered a bit. "A lovely dance." Then she had to laugh again. We had quite a bit of fun with everything. (i27, bereaved husband caring for his wife with a lung disease)*

Some caregivers found it hard to do nursing tasks such as washing because they had difficulties recognizing the naked and vulnerable patient as the partner or parent he or she used to be. For others, the relationship history guided their behavior, for example in taking care of a father who had always been dominant and harsh. Assertively *"switching the button"* from child role to caregiver role was helpful for some daughters to be able to do these tasks. While managing all kinds of practical caring activities could be difficult and tiring, for some, the burden of care lay in relational aspects:

> *I can discuss that with my husband, but sometimes you just notice – oh, what am I talking about? [...] I think, you should see, he looks so sick now, but I start talking about it [son in puberty] who once again did not get up [out of bed], yeah. Those are just normal things that are quite difficult for me. More difficult than cleaning or putting out the garbage cans. (i16, wife caring for her husband with cancer)*

An assertive attitude would also be demanded when several family members jointly provided care for their parent, and conflicts arose: what is best for our mother? Family caregiving could cause fights among siblings, with every child having an individual relationship with the parent while balancing their private life and needs. Sometimes, disappointments arose about the lack of involvement of other family members.

A last possibly tensed front requiring assertiveness was having a (volunteer) job while providing family care. Working and having contact with colleagues could be helpful, but some caregivers were not able to concentrate on tasks as before. Taking time off or receiving caring leave could be difficult. Some caregivers felt that they had to justify the fact they wanted to take sick leave to care for the patient themselves, instead of asking a neighbor:

> *I certainly felt abandoned when I had to defend myself before I could be present at the chemo days. I got remarks like "Do you really have to be there?" "Can't a neighbor or a good friend do that?" [...] It is, of course, completely ridiculous to say to someone who has just had very bad news, "Ask if the neighbor wants to go with her". (i12, bereaved husband caring for his wife with cancer)*

Overall, some interviewees wondered what it would be like for other caregivers, especially the less assertive ones or people with a migrant background:

> *…I am outspoken, I speak the language well, I use the medical terms correctly; and even then I don't get anywhere. How about the others who are not from the western world. [...] There are standards that you have to meet. If you don't fit, you are left out. (i31, bereaved daughter caring for her father with cancer)*

Discussion

This article presented a qualitative analysis of experiences with palliative family caregiving in Dutch home settings, focusing on the role of normativity. In sum, we first showed by whom caregivers feel *called to provide family care*, this study's central theme. Then we showed why people feel called, emphasizing how love, duty, or family dynamics motivate caregivers and in what way the

palliative care context intensifies feeling called upon with a sense of urgency. These motives, lastly, explain to what people feel called.

The limitlessness of feeling called to care

Feeling called to care starts not only with the obvious cause, e.g. a patient's need for care as the Informal Care Model [18] suggests, but can also be instigated by health care professionals entering a caregiver's normal life, well-intending family and friends, or it can be self-imposed. We will elaborate on this self-imposed call in the next section. Concerning the call instigated by professionals: previous research shows how professional home care provided security, allowing caregivers and patients to focus more on their family life and prepare for death [30], and offered relief when the right assistance was given on the right time [31]. Our findings confirm this, but also show how professionals entering normal life and routines can instigate a call towards family caregivers, in line with the finding that support sometimes overwhelms or adds responsibilities [31]. Our study also reinforces the idea that previous negative experiences with the healthcare system are motivating the provision of care at home [31], thus showing how caring practices are related to their sociopolitical context. Furthermore, concerning the call instigated by family and friends: caregivers' often close relationship to the patients and their gatekeeper role places them in a complex social web, highlighting the challenge of maintaining relationships during and beyond the dying process [32], in line with what we found to be a call by well-intending but sometimes demanding friends and family surrounding the caregiver and patient.

Although our results show that caregivers often feel constantly called upon, it should be emphasized that this does not necessarily equal feeling burdened. Caregivers often express ambivalent feelings: caring can be a positive, rewarding, or honoring experience, *while* at the same time feeling occupied or exhausted. In the light of sharing last moments and activities, specifically, bereaved caregivers in our study experienced gratitude for the time spent with their partner or family member, and for having been able to persevere and to give everything they could. Previous research also suggests that although feeling burdened by caregiving in a palliative care context, caregiving can help appreciating 'the little things' and becoming closer to the patient [8], and may be a personally meaningful and transforming experience [33].

From each of the found appearances of feeling called, it becomes clear that caring situations are, as care ethicists argue, *"situations that are given to us, that we find ourselves in, as a consequence of being related"* (p. 527) [34]. The families in this study often found themselves embedded in practices in which they felt called and responsible to care [19]. Our results suggest that – rather than in feeling called per se – the burden of care might lie in the seeming *limitlessness* to which people feel called, i.e. the self-ignorance that lies in being constantly alert and available, attentive to the patient first, and assertive in managing the situation while also facing dilemmas in balancing care with one's other needs. This resonates with the previously investigated all-consuming nature of caring in a situation of serious illness, overwhelming people with its demands [35], *"being on 24/7"* (p. 1232) [31], and with the observation that family care has an impact on a caregiver's whole personal realm, i.e. feeling the physical and emotional burden of home care and experiencing limitations in living normal daily life [16].

Previous research used existential psychology to understand the psychological complexity of caregivers' emotional challenges [36]. Our study, however, adds a further interpretation of caregiver responses to the felt calls, as they seem to be reinforced by the *normativity* essential to the actual caring situation. We suggest that it is the inherent normative nature of palliative family caregiving that invites us to rethink our concept of burden and to challenge our bias towards the negative aspects of caregiving [37, 38].

Normativity and social expectations in a palliative care context

The specific, e.g. palliative, care context seems crucial for understanding caregiver experiences, as it changes relationships [16] and intensifies the felt calls with a sense of urgency of having only one final chance to act. The assertive response to the felt calls we found resonates with findings on taking charge in coordinating home care and making important decisions, in which bereaved caregivers often *"felt thrust into this role without adequate recourses to fulfill its expectations"* (p. 1232) [39]. This reference to expectations underlines how normativity shapes caregiver experiences.

In addition to this specific context, the general normative aspects of love, duty, and family dynamics appeared to be a motivating force behind what caregivers expect of themselves. Interestingly, in family caregiving for people with dementia, the same aspects were found to be interwoven, paralleling the long-term and fluid caregiving role in that context [40]. That long-term care role was found in our study among participants caring for someone with end-stage chronic obstructive pulmonary disease, as opposed to a role that was more suddenly imposed and more directed towards an imminent death among participants caring for someone with metastasized cancer. Previous studies in palliative care also implicitly referred to normativity, for example in caregivers being determined

to care for their loved one at home out of *"love, respect, obligation, or giving back to someone who had given them so much and as a way to honor their ill family member's wishes"* (p. 1232) [31]. This article, however, focused on this normativity explicitly, aiming at providing a more complete overview of the normative aspects motivating caregivers.

A sociological perspective helps deepen our understanding of how these normative aspects are shaping caregiver experiences. Ultimately, family caregiving is an evolving experience, subject to social scripts and expectations [32]. According to sociologist Hochschild's framework about how people make sense of their emotions, we all live by 'framing rules' that govern how we view our situation, and by 'feeling rules' with which we relate to these frames and define what we should or should not feel [41, 42]. Caregivers, seen from this perspective, judge their experiences and feelings based on how well they fit with what they believe is to be expected from a 'good caregiver' [32]. Misfits inevitably lead to feelings of failure, Broom et al. (2019) argue: caregivers' actual but perhaps 'inappropriate' feelings then conflict with their intentions and, sometimes even romanticized, expectations, such as 'till death to do us part'. To a certain extent, these sociological mechanisms showed up in our study. Caregivers used such expressions as 'till death do us part' or 'for better or worse' as a motive to persevere while feeling exhausted, or they wanted to keep doing everything they could to prevent themselves from not feeling guilty afterward (the feeling of failure in Broom's terms). Or they regretted not having been able to be present when their family member died, showing the *"inevitable messiness of the dying process"* (p. 7) which obstructs caregivers' abilities to achieve their desired outcomes [32].

However, notably, we did not observe many explicit social scripts or expectations in our participant's stories. In comparison with the other three manifestations of feeling called, the fourth – the self-imposed call, stemming from one's convictions and values, seemed less thick. Perhaps, participants could or did not articulate these social expectations by themselves. This can be explained by the fact that our interview questions were focused on *personal* experiences and more on the relationship between caregiver and patient while less on the phenomenon of caregiving in general. Our results might imply that the interviewed caregivers were motivated purely intrinsically, as their limitless caring behavior would also suggest; or that these caregivers did not recognize the social expectations as motivating. This would imply that social scripts and pressure play a more implicit and subtle role.

Further research is needed to investigate a possible difference between reacting to a felt call out of intrinsic motives or because of what is expected. Future research specifically focused on normative aspects might benefit from a serial exploration. Interviewing the same caregivers serially, both before the death of the patient and during bereavement, might allow for showing ambivalent feelings that are suppressed in the dying process and can only be revealed in the bereavement phase [32].

Practical implications: rethinking caregiver support

Studies sometimes show average or median hours of family caregiving per week to indicate burden [11], but invested time is not the only issue nor the most relevant. Caregiving is complex, as we have shown by providing more insight into the inherent normative nature of family caregiving, and is also deeply connected to one's social, cultural, and political context and related feelings of power [19]. Revealing this complexity to caregivers gives counterweight to dominant expectations [32], and may enable them to talk about their seemingly inappropriate experiences or feelings – whether negative or positive. This article's insights also provide us with relevant clues about how to better understand and support caregivers.

As our study showed, caregivers want to stay close to the person they care for, reinforced by normative aspects. In our belief, the related pressure will not likely be resolved by professional home care or respite care. Respite might offer effective relief, provided that it is adjusted to caregiver needs, for example concerning confidence about the patient being in good hands [31]. However, reduction of caring hours or organizing activities outside the caregiver's home to take time-off might not be the only suitable solution for people who do not wish, dare, or feel able to leave their loved one. It was already suggested that being both relative and family caregiver might create a reluctance in asking for professional help, due to the dynamics of these sometimes conflicting dual roles [15]. Our results help us understand this reluctance. Feeling called upon or even burdened does not necessarily mean that caregivers do not want to provide care anymore, or that they wish support or respite. Taking into account the always-existing normative and complex nature of caregiving, support should not be aimed at liberating caregivers *from* the situation but supporting them *in* whatever overwhelms them in providing family care.

This implies a responsibility for healthcare professionals regarding the already widely acknowledged need for better collaboration between family caregivers and supporting professionals in the palliative home setting [9]. They should recognize that caregivers might be determined to enable their partner or family member to die at home, regardless of whether this is burdensome, and align their support with caregivers' own goals in caring at home [31]. In addition, support should enable caregivers

to pause and reflect on (and perhaps break) the social rules implicitly guiding their sometimes automatic responses concerning what is expected or appropriate [41]. Although we acknowledge that caring might be something families find themselves in (*"you just do that for each other"*, as an interviewee stated) and a caregiver is never a freely choosing person [19], we believe there is some degree of agency in how one provides family care. As our analysis showed a difference between feeling called and how people respond, we suggest that caregivers should be enabled to explore the feeling of being called upon – by whom, to what, and why –, to make them aware that, to some extent, they *do* have choices in how to respond. Such exploration helps to address their immediate and sometimes limitless caregiving behavior but also allows for tailoring support to individuals' needs. After all, family caregiving, regardless of being burdensome, can be an integral, and for some highly important and rewarding part of the relationship between partners or family members.

Strengths and limitations

As many studies focus on either current or bereaved caregivers [32], a strength of our study is that both were interviewed. In the bereaved group, however, it was difficult to distinguish between grief and the impact of the withdrawal of caregiving. The fact that we also interviewed patients has broadened our analysis. Another strength is that our analysis was conducted cyclically, and co-occurred with collecting the data.

A limitation is that it is likely that we have only interviewed people who felt able to talk about their experiences and thus consented to participate. We cannot rule out the possibility that certain caregivers were excluded from our sample, for example severely overburdened caregivers, or partners with a non-loving relationship affecting one's caregiving behavior and feelings. Theoretical sampling of such participants, although important for a grounded theory, turned out to be practically difficult, due to the unsteady and precarious circumstances of the study population, time constraints, and general difficulties in finding participants. Theoretical sampling with regard to the interviewees' personal circumstances (such as being deeply religious or having a troublesome relationship) was hindered by practical difficulties and privacy related issues. It was also difficult to reach interviewees with migrant backgrounds. The ratio of daughter/son in the sample of children who cared for a parent was relatively uneven as well, but at the same time reflects the dynamic of gender in care provision, as recognized in palliative care for older adults [43] or among people with a migrant background [41]. A further limitation in this regard is that we did not systematically document demographic variables of each participant, such as religion or social background, but left it to the initiative of the participants whether or not to share such information.

Furthermore, the interviews likely prompted participants to give words to their experiences, as interview research often provides people with an opportunity to reflect on their situation [44]. This can be regarded as both a strength and a limitation. Normative aspects especially are often implicit until made explicit by ethical reflection or questions; our participants might not have put this normativity into words by themselves. In addition, it can be hard for people to articulate social expectations [32] or express certain 'background worries' occupying their minds in light of more pressing matters in the palliative care situation [45]. A strength, in this respect, was our background as medical ethicists, leading to sensitiveness to normative aspects in caregivers' daily speech, although we did not specifically ask interviewees about this.

Conclusions

By providing more insight into the inherent normative nature of family caregiving of seriously ill patients within a palliative care context, we have added complexity and depth to our understanding of the much-investigated concept of caregiver burden. Our study indicates that feeling called to care, as far as it is burdensome, cannot – and perhaps should not – be resolved. Caring might be something families just find themselves in due to being related. Rather than in feeling called upon per se, we believe the burden of care lies in the seeming limitlessness to which people feel called. Social expectations play an important, yet often subtle and implicit, role here. Support, then, should enable people to explore the feeling of being called upon – by whom, to do what, and because of which norms or expectations. This helps to tailor support to individuals' needs. Family caregiving, regardless of being burdensome, can be an integral and for some highly important and rewarding part of the relationship between partners or family members.

Acknowledgments

We are grateful to the people who participated in this study for sharing their stories. We also want to thank the healthcare professionals who helped us to include interviewees. A special thanks to prof. Marianne Boenink for her constructive comments on an earlier version of this article.

Authors' contributions

JvG and GO developed the conceptualization of the study. All authors contributed to the methodology. MH collected and analyzed the data, JvG analyzed three interview transcripts independently as co-coder. MH was the main contributor in writing the manuscript, JvG and GO contributed to MH's analysis in ongoing peer discussions and participated in the preparation of the manuscript. All authors read and approved the final manuscript.

Funding
This study was funded by the Netherlands Organisation for Health Research and Development (ZonMw, dossier number: 844001310). The funding source had no involvement in the design, nor in the collection, analysis, and interpretation of data, nor in writing the manuscript.

Availability of data and materials
The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
Ethical approval was sought from the Medical Review Ethics Committee region Arnhem-Nijmegen (registration number 2017–3415), who determined that this study does not fall under the scope of the Medical Research Involving Human Subjects Act (WMO). All methods were carried out in accordance with the relevant guidelines and regulations. All interviewees gave their written informed consent to participate.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Received: 5 August 2021 Accepted: 13 October 2021
Published online: 27 November 2021

References

1. Choi S, Seo J. Analysis of caregiver burden in palliative care: an integrated review. Nurs Forum. 2019:1–11.
2. Alam S, Hannon B, Zimmermann C. Palliative Care for Family Caregivers. J Clin Oncol. 2020;38(9):926–36.
3. De Boer A, Plaisier I, De Klerk M. Mantelzorgers in het vizier. Beleidssignalement mantelzorg. Sociaal en Cultureel Planbureau: Den Haag; 2019.
4. De Klerk M, De Boer A, Plaisier I, Schyns P. Voor elkaar? Stand van de informele hulp in 2016. Den Haag: Sociaal en Cultureel Planbureau; 2017.
5. Plothner M, Schmidt K, de Jong L, Zeidler J, Damm K. Needs and preferences of informal caregivers regarding outpatient care for the elderly: a systematic literature review. BMC Geriatr. 2019;19(1):82.
6. De Boer A, De Klerk M, Verbeek-Oudijk D, Plaisier I. Blijvende bron van zorg. Ontwikkelingen in het geven van informele hulp 2014–2019. Sociaal en Cultureel Planbureau: Den Haag; 2020.
7. Morris SM, King C, Turner M, Payne S. Family carers providing support to a person dying in the home setting: a narrative literature review. Palliat Med. 2015;29(6):487–95.
8. Bijnsdorp FM, van der Beek AJ, Pasman HRW, Boot CRL, de Boer AH, Plaisier I, et al. Home care for terminally ill patients: the experiences of family caregivers with and without paid work. BMJ Support Palliat Care 2019;0:1–10.
9. Vermorgen M, Vandenbogaerde I, Van Audenhove C, Hudson P, Deliens L, Cohen J, et al. Are family carers part of the care team providing end-of-life care? A qualitative interview study on the collaboration between family and professional carers. Palliat Med. 2021;35(1):109–19.
10. Gomes B, Calanzani N, Gysels M, Hall S, Higginson IJ. Heterogeneity and changes in preferences for dying at home: a systematic review. BMC Palliative Care. 2013;12(1):7.
11. Rowland C, Hanratty B, Pilling M, Van den Berg B, Grande G. The contributions of family care-givers at end of life: a national post-bereavement census survey of cancer carers' hours of care and expenditures. Palliat Med. 2017;31(4):346–55.
12. Ahn S, Romo RD, Campbell CL. A systematic review of interventions for family caregivers who care for patients with advanced cancer at home. Patient Educ Couns. 2020;103(8):1518–30.
13. Oechsle K, Ullrich A, Marx G, Benze G, Heine J, Dickel LM, et al. Psychological burden in family caregivers of patients with advanced cancer at initiation of specialist inpatient palliative care. BMC Palliative Care. 2019;18(1):102.
14. Tarberg AS, Kvangarsnes M, Hole T, Thronæs M, Madssen TS, Landstad BJ. Silent voices: family caregivers' narratives of involvement in palliative care. Nursing Open. 2019;6(4):1446–54.
15. Funk L, Stajduhar K, Toye C, Aoun S, Grande G, Todd C. Part 2: home-based family caregiving at the end of life: a comprehensive review of published qualitative research (1998-2008). Palliat Med. 2010;24(6):594–607.
16. Martín JM, Olano-Lizarraga M, Saracíbar-Razquin M. The experience of family caregivers caring for a terminal patient at home: a research review. Int J Nurs Stud. 2016;64:1–12.
17. Van Roij J, Brom L, Youssef-El Soud M, Van de Poll-Franse L, Raijmakers NJH. Social consequences of advanced cancer in patients and their informal caregivers: a qualitative study. Support Care Cancer. 2019;27(4):1187–95.
18. Broese van Groenou MIB, De Boer A. Providing informal care in a changing society. Eur J Ageing. 2016;13(3):271–9.
19. van Nistelrooij I, Visse M. Me? The invisible call of responsibility and its promise for care ethics: a phenomenological view. Med Health Care Philos. 2019;22(2):275–85.
20. Randall FD, R. S. The philosophy of palliative care: critique and reconstruction. Oxford: University Press; 2010.
21. Kristanti MS, Effendy C, Utarini A, Vernooij-Dassen M, Engels Y. The experience of family caregivers of patients with cancer in an Asian country: a grounded theory approach. Palliat Med. 2019;33(6):676–84.
22. Woodman C, Baillie J, Sivell S. The preferences and perspectives of family caregivers towards place of care for their relatives at the end-of-life. A systematic review and thematic synthesis of the qualitative evidence. BMJ Support Palliat Care. 2016;6(4):418–29.
23. Rieger KL. Discriminating among grounded theory approaches. Nurs Inq. 2019;26(1):e12261.
24. Corbin J, Strauss A. Basics of qualititative research. Techniques and procedures for developing grounded theory. 4 ed. Los Angeles: SAGE Publications, Inc.; 2015.
25. Charmaz K. Constructing grounded theory. A practical guide through qualitative analysis: SAGE Publications Ltd; 2006.
26. Holton JA. The coding process and its challenges. In: The SAGE Handbook of Grounded Theory. London: SAGE Publications Ltd; 2007. p. 265–89. https://doi.org/10.4135/9781848607941.n13
27. Cornish F, Gillespie A, Zittoun T. Collaborative analysis of qualitative data. In: The SAGE Handbook of Qualitative Data Analysis; 2013: 79–93.
28. Hall WA, Long B, Bermback N, Jordan S, Patterson K. Qualitative teamwork issues and strategies: coordination through mutual adjustment. Qual Health Res. 2005;15(3):394–410.
29. Lempert LB. Asking questions of the data: Memo writing in the grounded theory tradition. In: The SAGE Handbook of Grounded Theory. London: SAGE Publications Ltd; 2007. p. 245–64. https://doi.org/10.4135/9781848607941.n12
30. Sarmento VP, Gysels M, Higginson IJ, Gomes B. Home palliative care works: but how? A meta-ethnography of the experiences of patients and family caregivers. BMJ Support Palliat Care 2017;7(4):0.
31. Robinson CA, Bottorff JL, McFee E, Bissell LJ, Fyles G. Caring at home until death: enabled determination. Support Care Cancer. 2017;25(4):1229–36.
32. Broom A, Parker RB, Kenny K. Authenticity, ambivalence and recognition in caring at the end of life and beyond. Soc Sci Med. 2019;239:112554.
33. Vachon M. "It made me more human": Existential journeys of family caregivers from prognosis notification until after the death of a loved one. J Palliat Med. 2020;23(12):1613-8. https://doi.org/10.1089/jpm.2019.0689.
34. van Nistelrooij I. Self-sacrifice and self-affirmation within care-giving. Med Health Care Philos. 2014;17(4):519–28.
35. Breen LJ, Aoun SM, O'Connor M, Howting D, Halkett GKB. Family caregivers' preparations for death: a qualitative analysis. J Pain Symptom Manag. 2018;55(6):1473–9.
36. Totman J, Pistrang N, Smith S, Hennessey S, Martin J. You only have one chance to get it right': a qualitative study of relatives' experiences of caring at home for a family member with terminal cancer. Palliat Med. 2015;29(6):496–507.

37. Ates G, Ebenau AF, Busa C, Csikos A, Hasselaar J, Jaspers B, et al. "Never at ease" - family carers within integrated palliative care: a multinational, mixed method study. BMC Palliat Care. 2018;17(1):39.
38. Van Groenou MIB, De Boer A, Iedema J. Positive and negative evaluation of caregiving among three different types of informal care relationships. Eur J Ageing. 2013;10(4):301–11.
39. Mohammed S, Swami N, Pope A, Rodin G, Hannon B, Nissim R, et al. "I didn't want to be in charge and yet I was": bereaved caregivers' accounts of providing home care for family members with advanced cancer. Psychooncology. 2018;27(4):1229–36.
40. Greenwood N, Smith R. Motivations for being informal carers of people living with dementia: a systematic review of qualitative literature. BMC Geriatr. 2019;19(1):169.
41. Ahmad M, Van den Broeke J, Saharso S, Tonkens E. Persons with a migration background caring for a family member with dementia: challenges to shared care. The Gerontologist. 2019;60(2):340–9.
42. Tonkens E, Van den Broeke J, Hoijtink M. In search of resonating pleasure. Cooperation between care workers, carers and volunteers and clients in long term care. [paper presentation] Who cares? Informal care as shock absorber of welfare state reform, Amsterdam School of Social Sciences Research, University of Amsterdam. 2008.
43. Morgan T, Ann Williams L, Trussardi G, Gott M. Gender and family caregiving at the end-of-life in the context of old age: a systematic review. Palliat Med. 2016;30(7):616–24.
44. Aoun S, Slatyer S, Deas K, Nekolaichuk C. Family caregiver participation in palliative care research: challenging the myth. J Pain Symptom Manag. 2017;53(5):851–61.
45. Thomas C, Turner M, Payne S, Milligan C, Brearley S, Seamark D, et al. Family carers' experiences of coping with the deaths of adults in home settings: a narrative analysis of carers' relevant background worries. Palliat Med. 2018;32(5):950–9.

Publisher's Note

Erkenntnis
https://doi.org/10.1007/s10670-021-00430-9

ORIGINAL RESEARCH

Epistemic Blame and the Normativity of Evidence

Sebastian Schmidt[1]

Received: 1 November 2020 / Accepted: 1 June 2021

Abstract
The normative force of evidence can seem puzzling. It seems that having conclusive evidence for a proposition does not, by itself, make it true that one ought to believe the proposition. But spelling out the condition that evidence must meet in order to provide us with genuine normative reasons for belief seems to lead us into a dilemma: the condition either fails to explain the normative *significance* of epistemic reasons or it renders the *content* of epistemic norms practical. The first aim of this paper is to spell out this challenge for the normativity of evidence. I argue that the challenge rests on a plausible assumption about the conceptual connection between normative reasons and blameworthiness. The second aim of the paper is to show how we can meet the challenge by spelling out a concept of epistemic blameworthiness. Drawing on recent accounts of doxastic responsibility and epistemic blame, I suggest that the normativity of evidence is revealed in our practice of suspending epistemic trust in response to impaired epistemic relationships. Recognizing suspension of trust as a form of epistemic blame allows us to make sense of a *purely* epistemic kind of normativity the existence of which has recently been called into doubt by certain versions of pragmatism and instrumentalism.

1 Introduction

Do epistemic norms provide us with normative reasons for compliance? Such norms tell us, very roughly, that we should believe what we have sufficient evidence for, and that we should refrain from believing what we lack sufficient evidence for. Recently, epistemologists have questioned that epistemic norms have genuine normative significance. Susanna Rinard, for instance, argues "that only pragmatic considerations are genuine reasons for belief. That is, purely evidential considerations—evidential considerations that are not also pragmatic reasons—do not constitute reasons for belief" (2015, 219). To illustrate this view, consider a case in which you happen to

✉ Sebastian Schmidt
sebastian.schmidt@philos.uzh.ch

[1] Philosophisches Seminar, University of Zurich (UZH), Room G-228, Zürichbergstrasse 43, 8044 Zurich, Switzerland

Published online: 14 June 2021

come across the newest celebrity gossip in a magazine that you know to be reliable. Assume that you know that having a true belief about the gossip is not, nor will ever be, of any practical value for you, and that you are not curious about such gossip at all. Now consider the following questions:

- *Should* you believe the gossip?
- Are you *blameworthy* or *criticizable* if you fail to believe the gossip?

Or consider a slightly altered case in which you happen to *believe* some celebrity gossip although you are aware that the magazine in which you came across it is *not* reliable. Your belief is not based on sufficient evidence. We can again stipulate that you know that nothing bad will ever come from having such an ill-based belief about this unimportant matter that is of no interest to you. Now consider the following question:

- Are you *blameworthy* or *criticizable* for having this belief?

The view under consideration in this paper would reply with "no" to all these questions. Let us call this view "anti-normativism about evidence" (short: ANE). Proponents of ANE argue that mere evidence does never provide us, *by itself*, with a (normative) reason for belief. Next to Rinard's view, recent instrumentalists about reasons for belief commit to ANE. Asbjørn Steglich-Petersen and Mattias Skipper argue "that evidence for *p* speaks in favor of believing *p* only in context where there is a practical reason to pursue the aim of coming to a true belief as to whether *p*" (2019, 9), and that therefore "it is strictly speaking false to say that evidence *by itself* constitutes a normative reason for belief" (2020, 114). Similarly, Maguire and Woods (2020) have recently denied that purely epistemic norms provide us with reasons.[1] They compare epistemic norms with rules of games: we only have a reason to comply with each if we have a practical (prudential or moral) reason to engage in the relevant practice. That is, I have a reason to move a chess piece according to the rules only if I have a practical reason to play chess; analogously, they argue that I have a reason to believe that p only if I have a practical reason to play what Maguire and Woods call "the game of belief".[2]

[1] In their terminology, epistemic norms are just not *authoritatively* normative. However, according to their view, also the rules of chess count as (non-authoritatively) normative. In my terminology, norms of chess are *not* normative in the sense that they do not provide us with reasons for compliance. I follow Kiesewetter (2017, 3–4) in this use of "normative". In Maguire's/Wood's terminology, I am interested in *authoritative* norms and reasons. While I use the term "norm" as including also non-normative rules or standards, I use "reasons" exclusively in a normative sense. I do not think there are non-normative reasons (except for explanatory ones). So my question is: "Do epistemic norms provide us with reasons?" Cf. Hofmann (2020) on some unpacking of the idea that some norms "provide" us with reasons—on which I cannot elaborate within the scope of this paper.

[2] For recent accounts with similar implications, cf. Bondy (2018), Cowie (2014), Mantel (2019), McCormick (2015, 2020), Papineau (2013). Of these, only McCormick identifies explicitly as a pragmatist (like Rinard). Instrumentalists argue that they can avoid a commitment to pragmatism—i.e., to the existence of practical reasons for belief per se—by saying that it is always *evidence* that motivates our beliefs (rather than practical reasons). Instrumentalists merely claim that evidence *gains its normative*

The normativity of evidence is at stake in contemporary epistemology. What could possibly count as a satisfying reply to ANE? This paper is devoted to answering this question. I argue that, in order to get clear about the normative significance of evidence, we need to think about the reactive attitudes that are appropriate towards violations of purely epistemic norms. I reach this claim by first spelling out the dilemma of explaining why epistemic norms matter without thereby rendering them practical (Sect. 2). This dilemma will allow us to pin down a challenge for normativism about evidence: finding a satisfying conception of a distinctively epistemic kind of blame (Sect. 3). I then propose, in outline, a reply to this challenge (Sect. 4). I suggest that we hold each other answerable to epistemic norms by showing reactive attitudes towards each other's epistemic failures—mainly suspension of epistemic trust (cf. Boult, 2020, 2021). The normativity of evidence can become intelligible by understanding this practice. For the reactive attitudes within our epistemic practice reveal the normative significance of purely evidential considerations.

The result of the paper is that we should not submit too quickly to treating the epistemic as hostage to the practical. Rather, we should first engage in the project of understanding the distinctive normative significance of purely evidential considerations by appealing to our responsibility for (non-)compliance with epistemic norms. Only if it were to turn out that this project fails—because, say, it turns out that there is no distinctive responsibility attached to the purely epistemic—we would be justified to endorse ANE. However, proponents of ANE have not yet provided arguments that this project fails. Doing so would require them to engage with recent accounts of epistemic blame and responsibility for belief. The paper thus shifts the dialectical burden to proponents of ANE and connects debates within contemporary epistemology.

2 The Challenge for the Normativity of Evidence

This section argues that normativism about evidence—the view that purely evidential considerations provide us with reasons for belief—faces a dilemma. I first outline two strategies for finding a plausible content of epistemic norms (Sect. 2.1). I argue that both strategies give rise to the same dilemma for normativism (Sects. 2.2). This will then allow us to formulate the central challenge for normativism in Sect. 3, and thus to see the intuitive appeal of ANE.

Footnote 2 (continued)

authority from practical considerations (cf. Cowie, 2014, 4004–5; Steglich-Petersen & Skipper, 2019, 11). I am not here interested in subtle differences between the versions of pragmatism and instrumentalism I consider. I am only interested in their implication of ANE. (However, cf. note 10 below on how Steglich-Petersen/Skipper's version of instrumentalism might collapse into pragmatism.).

2.1 Replying to the Clutter-Objection: Background Conditions on Epistemic Norms

Consider the following rough first approximation towards formulating an epistemic norm (also mentioned at the beginning of this paper):

> (EN) One ought to believe everything that is sufficiently supported by one's evidence.

Gilbert Harman (1986, 12) points out that (EN) implies that we should clutter our minds with uninteresting implications of our beliefs. My current belief-stock implies the proposition that [I am sitting in my office *or* the moon is made of cheese *or* there is no Corona virus *or* there is no human-induced climate change *or* …]. This disjunctive proposition is true right now while I am sitting in my office. It is true because the first claim, that I am sitting in my office, is true. At the same time, it seems that I do not always believe this disjunctive proposition while I am sitting in my office. Most importantly, it seems that I would not be *blameworthy* or *criticizable* in any sense for *not* believing such disjunctive propositions. It follows, so it seems, that there is no unconditional norm to believe everything that is sufficiently supported by my evidence.[3]

In reaction to this, we might modify the epistemic standard so that its violation more plausibly gives rise to serious criticism. We might propose background conditions for *when* we are required to believe what our evidence sufficiently supports. These background conditions should fulfill two criteria:

(a) They must make it plausible that the subject is, at least normally or in paradigm cases of an epistemic norm violation, blameworthy or criticizable for not complying with the epistemic norm when the background conditions are fulfilled.
(b) They should not render the norm *practical* rather than *epistemic*.

Call (a) the criterion of *significance*, and (b) the criterion of *content*. (b) makes sense as a criterion on epistemic norms for our purposes because the aim of the normativist is to preserve a purely epistemic kind of normativity. But why (a)?

The guiding idea behind (a) is that the significance of a norm expresses itself in the reactive attitudes that we show towards violations of the norm. For instance, the significance of a moral requirement will make it often—in absence of an excuse or exemption—appropriate to show resentment or indignation. These emotions are expressions of the normative significance we attach to the moral requirement

[3] One might doubt that I do not believe the disjunctive proposition. For if I was asked whether I believe it, and I understand the content of the proposition, I will reply that I do believe it. However, even if one thinks that we believe all those disjunctive propositions, one will agree that we do not believe all the implications of our current belief-stock, like certain mathematical or logical implications that are just too hard to figure out. That we do not believe those implications does not make us blameworthy or criticizable in any sense. Furthermore, we *can* imagine a case where I fail to believe such weird disjunctive propositions. Why on earth, we might ask, should any reasonable person care about this so much as to regard me as blameworthy for not believing them?

because they are appropriate in face of its violation. Similarly, if there are distinctively epistemic norms that provide us with reasons for compliance, then we should expect there to be distinctively epistemic reactive attitudes that we show towards the violation of those epistemic norms. In this vein, Antti Kauppinen understands genuine norms (as contrasted with mere evaluative standards) as "rules that someone is accountable for conforming to in suitable circumstances" (2018, 3).

Here is an argument for (a). *Why* is it false that we should clutter our minds with all the implications of our beliefs? If we would accept that epistemic norms require us to clutter our minds, then we would constantly violate an epistemic norm by not drawing all the implications from our beliefs. However, this constant violation would have no further significance: we would not normally be blameworthy or criticizable for failing to believe what we epistemically ought to believe. The problem with this is that the normative force of this "ought" would then be mysterious: why comply with this norm if one cannot hold us legitimately responsible for non-compliance? The norm would at best have the force of the norms of etiquette or the rules of a game: we can intelligibly ask *why* we have a reason to comply with the norms of etiquette or rules of a game in a given situation. Such norms do not, by themselves, provide us with reasons. Thus, the best explanation of the intuitive appeal of Harman's clutter-objection when it comes to trivial implications of our beliefs is that we assume that epistemic norms with normative significance would fulfill (a).[4]

I will return to the connection between epistemic reasons and epistemic blameworthiness in Sect. 3. For now, consider another strategy for finding a plausible norm of belief that is purely epistemic. Instead of proposing background conditions to (EN), we might rather argue that (EN) is not a central epistemic norm at all. In response to Harman's clutter-objection, we might argue that, although we are never blameworthy merely for failing to believe what our evidence sufficiently supports, there are *other* epistemic requirements that are purely evidential. Specifically, we might defend the following epistemic norm:

> (EN*) One ought not [to believe what is not sufficiently supported by one's evidence.]

(EN*) is not confronted with Harman's clutter objection: rather than requiring us to believe plenty of propositions we intuitively have no reason to believe, (EN*) merely *prohibits* us to have certain beliefs. Steglich-Petersen (2018) accepts (EN*) but denies (EN): he thinks that evidence alone determines the *permissibility* of belief (which beliefs I am epistemically allowed to have), but he argues that evidence alone never gives us, as he puts it, "positive reason" to believe a certain proposition. Epistemic norms, on this picture, determine the space of doxastic permissibility, but they

[4] For the sake of brevity, I use the notion "blameworthiness" in what follows to cover both the notion of blame and other forms of personal criticism (cf. Kiesewetter 2017, ch. 2 on the notion of *personal* criticism as contrasted with, say, criticism of a knife for not being sharp). One might think that "blameworthiness" is essentially a moral notion. As I will argue in Sect. 4, epistemic blame is a *sui generis* kind of blame. This might warrant the label "criticism" instead of "blame" for epistemic blame. However, I take this to be a mere terminological issue.

never require a specific belief—rather, they merely prohibit certain beliefs (cf. also Whiting, 2010, 2013).

However, appealing to a norm of permissibility like (EN*) instead of (EN) won't help to defend the normativity of evidence against the challenge I spell out in this paper. First, it seems that (EN*) should not be any more plausible to skeptics about the normativity of evidence than (EN). The norm that we ought *to believe everything* that is supported by our evidence faces the problem that it requires us to needlessly clutter our minds. The norm that we ought *not to believe anything* that is *not* sufficiently supported by our evidence faces the reverse problem: it would require us not to have a lot of beliefs we, it seems, have no reason to give up. For example, why should I give up evidentially unsupported but beneficial beliefs? We often overestimate our own abilities or the virtues of our significant others. Arguably, this can promote our self-esteem (cf. Kelly, 2003) or our relationships (cf. Stroud, 2006). It seems that such beliefs are blameless as well. A general norm not to believe what is insufficiently supported by one's evidence seems *too exclusive*. And a norm to believe anything that is sufficiently supported by one's evidence seems *too inclusive*.

Furthermore, cases of trivial belief pose the same problem for (EN*) as they pose for (EN). What if you believe, in absence of sufficient evidence, that the celebrity gossip in this unreliable magazine is true? Why should it make sense for anyone to blame or criticize you for this belief, if we stipulate that your trivial belief will have no bad consequences? Such trivial propositions seem to pose a challenge to (EN*) as they do to (EN)—as I will illustrate in some more detail in Sect. 2.2 below.

One final clarificatory remark: I will call any form of blame that arises from the violation of a *purely* epistemic norm—i.e., an evidential norm that does not mention any practical considerations—*epistemic blame*. That is, epistemic blame, if there is such a thing, is a kind of negative reaction that is appropriate towards violations of purely epistemic norms like (EN) or (EN*), given suitable non-pragmatic background conditions. I now turn to the idea of background conditions on epistemic norms in some more detail to spell out a dilemma for normativism about evidence that will give rise to a challenge for normativism.

2.2 Proposing Background Conditions: A Dilemma

One way of developing a background condition on (EN) that might allow us to preserve the normativity of evidence is presented by Kiesewetter (2017, 184–5). He responds to Harman's clutter-objection by proposing that the central standard of theoretical rationality is to believe what one's evidence sufficiently supports *if one attends to this evidence*. According to this proposal, if I attend to the fact that I have sufficient evidence for a specific disjunctive proposition, then I would be criticizable (because irrational) if I do not come to believe it. Thus, Kiesewetter concludes, there is a sense in which I *ought* to believe it as soon as I *attend* to my sufficient evidence. Analogously, we could propose a background condition on (EN*) by saying that if we attend to the fact that we lack sufficient evidence for p, we ought not to believe p: we would be criticizable if we were to believe p; but we wouldn't be criticizable for

this if we never noticed how our belief lacks evidential support—we wouldn't count as irrational.[5]

It seems that Kiesewetter's background condition, while doing a good job in fulfilling criterion (b), does not fulfill (a). There are cases where we attend to sufficient evidence but where it would not make much sense to regard us as blameworthy or criticizable if we, for whatever reason, do not believe what the evidence supports. Take, again, the case in which I come across the newest celebrity gossip in a magazine that I know to be reliable, but I fail to believe the gossip. Assume again that having a belief about the matter is of no importance and that I do not care about whether the gossip is true. Proponents of ANE will argue that there is then no sense in which I am blameworthy, and that it is false that I ought to believe the gossip.[6]

It thus seems that if it does not *matter* whether we believe an evidentially well-supported proposition, it is false that we *ought* to believe it. However, if we instead propose a background condition on sufficient evidence that implies that it always matters whether we comply with the epistemic norm, we seem to end up violating criterion (b): if the norm is only in place when it matters whether we comply with it, then, so it seems, the norm is no longer a purely epistemic norm. It thus seems that there is no background condition on purely epistemic norms like (EN) and (EN*) that fulfills both (a) and (b). The normativist is in a dilemma.

Let us provide this dilemma with additional support by considering a background condition that does not fulfill criterion (b). Steglich-Petersen (2011), after discussing a case of a trivial belief that the subject is not required to have although it is well-supported by the subject's evidence (23), presents the following partial analysis of reasons for belief:

> Necessarily, if S has all-things-considered reason to form a belief about *p*, then [if S has epistemic reason to believe that *p*, S ought to believe that *p*] (24).

Here "epistemic reason" can be read as "sufficient evidence for p". The conditional then states that *if one has an all-things-considered reason to form a belief about p*, one ought to believe what one's evidence sufficiently supports. The italicized if-clause is Steglich-Petersen's proposed background condition for the epistemic standard (EN). Steglich-Petersen could analogously propose a background condition on (EN*): *if one has an all-things-considered reason to form a belief about p*, then if p is *not* sufficiently supported by one's evidence, one ought not to believe p.[7]

[5] Kiesewetter is not concerned with background conditions on (EN*). But his view might naturally be extended to (EN*) in the way described. As argued at the end of the last section, a background condition on (EN*) is as important as a background condition on (EN) to make it intelligible that these norms have reason-providing force.

[6] Cf., e.g., Rinard (2015, 220), Steglich-Petersen (2011, 23) for this verdict about structurally analogous cases.

[7] It is important to note that Steglich-Petersen does not think that this background condition is necessary for (EN*) to express an epistemic requirement. However, as I have argued in Sect. 2.1, he thereby ignores that (EN*) is faced with very similar challenges as (EN): Why give up a lot of beneficial beliefs that are insufficiently supported by one's evidence? And why blame or criticize anyone for believing something that is insufficiently supported by their evidence if it doesn't matter at all whether they believe it? It

Steglich-Petersen's "all-things-considered reason to form a belief about p" can, for instance, be a reason for an action prior to the belief.[8] "Forming a belief about p" might refer to the action of *thinking about whether p*: I may have more or less reason to think about whether something is true. I have some reason to think about whether there will be nice weather during the next days, but I have no reason at all to think about the newest celebrity gossip (I might even have reason to avoid such thinking). Thus, according to one plausible reading, the truth of "S ought to believe that p" in Steglich-Petersen's analysis is conditional on a practical reason for an action. It says that if we have a reason to bring it about or to maintain that we have a (true)[9] belief about p, and there is sufficient evidence for p, then we ought to believe that p.

Since we only have a reason to bring a belief about when it *matters* whether we have this belief, Steglich-Petersen's proposal does a good job fulfilling our criterion of *significance*. Yet his proposed background condition, and thus the proposed epistemic norm, is no longer *epistemic*, because it includes a practical reason (for an action). His proposal thus fails to fulfill our criterion of *content*.[10]

Thus, while proposing non-pragmatic background conditions on epistemic norms (á la Kiesewetter) apparently does not result in doxastic norms that fulfill the criterion of *significance*, proposing a pragmatic background condition (á la Steglich-Petersen) results in doxastic norms that do not fulfill the criterion of *content*. If the normativist about evidence accepts the criterion of significance for epistemic norms, they have to defend the claim that compliance with epistemic norms matters (in a sense) even if we do not equip these norms with a pragmatic background condition. *Prima facie*, it is hard to see how purely epistemic norms could matter by themselves. Therefore, any normativist will, it seems, end up in either of two horns:

(i) Epistemic norms are purely epistemic, but they fail to be significant.
(ii) Epistemic norms are significant, but they fail to be purely epistemic.

Footnote 7 (continued)
seems that Steglich-Petersen's instrumentalist framework commits him to a background condition not only for (EN), but also to (EN*): without assuming pragmatic background conditions, both norms seem questionable as genuine normative requirements.

[8] Note that the reason for forming a belief about p cannot be itself an epistemic reason for belief, because epistemic reasons favor believing a specific proposition—they favor believing that p or not believing that p. Steglich-Petersen's reason to form a (true) belief about p, by contrast, does not favor believing a specific proposition. It merely favors having a true belief about a matter, whatever this belief turns out to be. Only practical reasons can favor forming a belief about a matter without favoring a specific proposition.

[9] In later works, Steglich-Petersen accepts that the reason to form a belief about p must in fact be a reason to form a *true* belief about p (see esp. the formulations of the norms of belief in Steglich-Petersen & Skipper, 2019, 2020).

[10] One might wish to interpret Steglich-Petersen's view without committing to the idea that the practical reason is always a reason for an action of managing one's beliefs (causing, maintaining, etc.). However, as pointed out in note 8 above, the "reason to form a belief about whether p" cannot be epistemic. It must thus be practical. If the practical reason does not favor an action, it must favor the state of believing itself. This would, however, commit Steglich-Petersen to pragmatism about reasons for belief. Whether this reason favors actions of managing beliefs or rather beliefs themselves, the condition renders the requirement practical rather than purely epistemic.

We might be tempted conclude from the dilemma that ANE is true: we might think that we should just reject the idea that we are ever *epistemically* blameworthy, and that there is any such thing as a purely *epistemic* kind of normativity.

I think this dilemma points to a serious challenge for the normativity of evidence. However, I will ultimately propose that purely evidential considerations *have* a kind of normative significance: (non-)compliance with purely epistemic norms matters. This requires me to make sense of a notion of epistemic blame. For the significance of a norm expresses itself in our reactive attitudes towards violations of the norm (cf. Sect. 2.1). Thus, giving a satisfying reply to ANE requires us to reconsider the concept of blameworthiness and responsibility for beliefs. For now, however, we should accept that there is a good case to be made for ANE. I now turn to the argument that is the core of this challenge for the normativity of evidence.

3 An Argument for Anti-normativism About Evidence

This section presents an argument that constitutes the core of the challenge for normativists (i.e., opponents of ANE) (Sect. 3.1), and it defends one of the two premises of this argument (Sect. 3.2). Replying to the challenge will thus require rejecting the other premise of the argument, which states that there is no distinctively *epistemic* kind of blame. I will turn to the rejection of this premise in Sect. 4.

3.1 The Argument from Doxastic Blameworthiness

Given the dilemma spelled out in Sect. 2, it is now easy to see how the denial of epistemic blame can give rise to an argument for ANE. Take "purely epistemic norms" to refer to (EN) or (EN*), and, if you want to, include any background conditions on purely epistemic norms that do not render the norm practical (like Kiesewetter's attending condition):

(1) Evidence provides us with epistemic reasons for belief only if we can be blameworthy (or criticizable)[11] for violating purely epistemic norms.
(2) We cannot be blameworthy (or criticizable) for violating purely epistemic norms.
(3) Thus, evidence does not provide us with epistemic reasons for belief.

According to the view expressed in (3), evidence provides us with reasons only if there is some practical value realized in following the evidence. Even if evidence does not *by itself* provide us with reasons for belief, it *appears* as if it provides us with such reason, because we normally have reason to follow the evidence—but this reason is only instrumental because it derives from our reason for the aim of gaining

[11] Cf. note 4 on my use of "blameworthy".

practically valuable true beliefs (cf. Rinard, 2015, 219). ANE just is this practical-instrumental approach to purely epistemic normativity.[12]

Premise (1) has a high *prima facie* plausibility. For it states a very minimal conceptual connection between reasons for belief and blameworthiness. The consequens of (1) states that it must be *possible* to be blameworthy for non-compliance with purely epistemic norms (under adequate non-pragmatic background conditions). (1) states that if this is not possible, evidence does not, by itself, provide us with reasons for belief. That is, (1) states that epistemic blame, as defined at the end of Sect. 2.1, must be conceptually possible if purely evidential considerations are to provide us with reasons for belief.

Remember that, as I have argued in Sect. 2.1, Harman's clutter-objection gets its grip on us only because we implicitly assume a connection between reasons and blameworthiness. Epistemologists propose background conditions on epistemic norms, like Kiesewetter's attending-condition or Steglich-Petersen's reason for forming belief about whether p, precisely because they want to make sense of the *significance* of these norms: they want to explain why it *matters* to us whether we comply with the norms—why we can be blamed or criticized if we fail to comply with them. The point of spelling out a notion of epistemic blame is to understand the normative force of evidence: why it matters to comply with epistemic norms.[13]

Yet most importantly, even if we were to reject that we need to be *always* blameworthy or criticizable for failing to comply with purely epistemic norms if evidence is to provide us with epistemic reasons for belief, this would not refute (1). For according to (1), it must merely be *possible* to be blameworthy for such non-compliance: there must be some possible cases in which we are blameworthy in virtue of the fact that we violate purely evidential norms in order for evidence to provide us with epistemic reasons for belief.

If (1) is indeed such a weak claim, then we should expect (2) to be more controversial to allow for the controversial conclusion ANE. Yet we saw that there is a strong *prima facie* case for (2) to be made by appealing to cases of trivial belief. It is not straightforward in what sense a person who violates a purely epistemic norm is blameworthy when nothing of practical value hinges on whether the person complies with the norm. That we *can* be blameworthy for violating a purely epistemic norm—i.e., non-(2)—seems to imply that we sometimes *are* blameworthy in cases of trivial belief. For cases of trivial belief are the cases in which we have isolated any non-epistemic factors—like the factor that there is a practical reason to consider our evidence carefully, or the factor that that it would be morally good to believe what one's evidence supports, or that having a certain belief would be disrespectful

[12] Cf. note 2 for proponents.

[13] A recent account of epistemic blame which is clearly motivated in this way is Kauppinen (2018). On the side of the proponents of ANE, McCormick (2020) argues that all blame for belief is ultimately prudential or moral. She *therefore* rejects that there is a normative domain of the purely epistemic. Thus, she explicitly reaches ANE by accepting (1) and (2) (cf. McCormick, 2020, 30). The problem with McCormick's argument is that she does not consider the epistemic forms of blame that have been spelled out by Boult (2020, 2021), Brown (2020), and Kauppinen (2018), and to which I will appeal in Sect. 4.2. She rather thinks of blame mainly in moral terms.

to a person. If there is never any blameworthiness left after we have isolated these non-epistemic factors—as it seems to be the case in cases of trivial belief—then it seems that there is no such thing as a distinctively epistemic kind of blameworthiness, and thus, given (1), no purely epistemic normativity. (I return to how normativists about evidence can deal with cases of trivial belief in Sect. 4.3.)

I will first defend (1) against objections. I thereby show that a challenge for normativism about evidence consists in arguing against (2)—i.e., in finding a good account of the normative significance of evidence by appealing to the concept of epistemic blameworthiness.

3.2 Some Worries About Blameworthiness as a Precondition on Reasons

The first objection to premise (1) points out that children or some non-human animals can act for reasons but cannot be blameworthy if they fail to do what is decisively supported by their reasons, because they are not responsible agents. Analogously, some children and animals might be considered has having reasons for belief even though they cannot be blameworthy for violating purely epistemic norms. Thus, (1) does not seem to hold.

The objection can easily be met by pointing out that (1) does not imply that *everyone* can be blameworthy for violating purely epistemic norms. As I have explained in the last subsection, (1) merely states that it must be possible to be blameworthy in a distinctively epistemic sense if evidence is to provide us with epistemic reasons for belief. If there are some beings who are not fully responsible agents or have not yet developed to fully responsible agents, they might also be exempted from epistemic blame. But that does not count against the idea that those beings who are fully responsible for their actions and beliefs must sometimes be subject to epistemic blame if evidence is to provide us with epistemic reasons.

Secondly, one might wish to deny (1) if one is an objectivist about the meaning of "ought" and "reasons". Objectivists deny a close connection between failing to do what one ought to and being blameworthy (a connection usually utilized or argued for by subjectivists).[14] Objectivism about practical reasons states that "S ought to φ" means that φing is the best option, no matter whether S is in a position to know or has some kind of cognitive access to the fact that φing is the best option. For example, if my house is burning even though I do not have any clue that it is burning, the objectivist would claim that I ought to leave the house. The subjectivist would deny this and say that I only *ought* to leave the house if I—in some way or another—have cognitive access to the fact that the house is burning. If I have access to that fact but do not leave the house, then I am blameworthy. According to the objectivist, it could be the case that I am not blameworthy, and yet it is true that I *ought* to leave the house (namely, if I am not in a position to know that it is burning). It thus seems that, if we are objectivists, we do not think that there is any close connection

[14] For some discussion and an argument for subjectivism that builds on the notion of *praise*worthiness, cf. Lord (2018, ch. 8). See also Kiesewetter (2017, ch. 8) for a good overview of the debate and another case for subjectivism.

between what we ought to do or what we have reason to do, on the one hand, and blameworthiness, on the other. This suggests that, *prima facie*, adopting an objectivist theory about reasons for belief would pose a problem to (1).

However, (1) is uncontroversial for objectivists. This is so again because (1) states a very *loose* connection between reasons and blameworthiness. Even if we grant the objectivist that we can be completely ignorant of our reasons, there will at least be some possible cases in which we are blameworthy for failing to give the response that is best. When we focus on actions rather than beliefs, such cases will be cases where we either act against our better knowledge of what is best or where we culpably fail to know what is best and do the wrong thing as a result of our ignorance. Analogously, if we are objectivists about reasons for belief—for instance, if we think that we ought believe only what is true or correct to believe rather than what is supported by our evidence –, then we can still argue that we are at least *sometimes* blameworthy because we have a false or incorrect belief. We would be blameworthy in some of the cases where the evidence was accessible to us, and yet did not form the correct doxastic attitude in accord with our evidence. So even if we spell out purely epistemic norms in objectivist terms (e.g., "one ought to believe only what is true"), this does not yet give us an argument against (1). For the accessibility of the evidence does not render the epistemic norm practical, and thus does not compromise the idea that evidence is normative only if we are *sometimes* blameworthy for violating purely epistemic norms.

Finally, one might want to object to (1) by adopting a permissivist epistemology. Permissivism states, roughly, that our total set of evidence permits more than one set of doxastic attitudes to take towards each (or at least some) proposition(s).[15] According to a permissivist, it could be true that if we have sufficient evidence for a proposition, it is both epistemically permissible to believe it as well as epistemically permissible not to believe it. Such an account might seem to be exactly the conclusion to draw from Harman's clutter-objection (cf. Sect. 2.1): we are not rationally *obligated* to believe anything that our evidence sufficiently supports, and thus we are not blameworthy for not drawing all implications from our beliefs. But it is always *permissible* for us to believe propositions that are sufficiently supported by our evidence.

However, it is not straightforward how permissivism could pose a problem for (1). The premise states that we can be blameworthy merely for violating a purely epistemic norm if evidence is to provide us with epistemic reasons. To deny this, the permissivist would have to argue that the possibility of epistemic blame is not a necessary condition on the normativity of evidence. They would have to claim that norms of epistemic permissibility provide us with epistemic reasons even if we cannot be blameworthy for violating them.

But this seems false. If something is permitted only under a certain condition, then it is *not* permitted—and thus *prohibited*—if this condition is not fulfilled. That is, that I am permitted to believe that p only if p is sufficiently supported by my

[15] The denial of permissivism is often discussed as the *Uniqueness Thesis* (as introduced by Feldman, 2007). On epistemic permissivism and some of its problems, cf. White (2005).

evidence implies that I *ought not* to believe that p whenever p is *not* sufficiently supported by my evidence. This is what (EN*) states. Thus, even if we understand the normative force of evidence in terms of permissibility, this will still require us to make sense of the idea that we sometimes *ought* to have certain doxastic attitudes under certain conditions. For instance, when p is insufficiently supported by evidence, we would be required either to suspend judgment about p or to disbelieve p. Purely epistemic norms would then be those norms that require us to either suspend belief about propositions or to disbelieve those propositions that are not sufficiently supported by our evidence (given suitable non-pragmatic background conditions). The normative significance of these norms would still be puzzling if we could not be blameworthy for violating them. Thus, if one reformulates (EN) as a claim about permission rather than obligation, one thereby commits to the equally puzzling epistemic norm (EN*). This won't make the normative force of reasons for belief any more intelligible if one does not show how one can be blameworthy for violating such purely epistemic norms.

This confirms that (1) is uncontroversial, mainly because it rests on a very loose connection between epistemic reasons and epistemic blameworthiness. The premise can be accepted both by proponents and opponents of the normativity of evidence, by objectivists, and by permissivists alike, and is thus a hinge around which the debate can progress.

It is important to see, however, that (1) would be false if we assume that the relevant sense of "blameworthy" must be a paradigm form of *moral* blame. A person's epistemic failure does not, by itself, give rise to emotions like resentment, indignation, or guilt. If someone believes that candidate X will win the next elections because the flight of the birds gave them a sign, then our reactive attitudes are not, or not necessarily, of that moral kind. But this does not yet rule out that there could sometimes be a distinct kind of *epistemic* blame appropriate in such cases even when the moral reactions aren't appropriate. This is also why (2) is not trivial: it claims that there is no such distinctively epistemic kind of blame.

I will now turn to the form of epistemic blame that can legitimately arise due to our epistemic answerability, and to my proposal of how this concept of epistemic blame might help us to understand the normativity of evidence.

4 Making the Normativity of Evidence Intelligible: A Proposal

If premise (1) is right—if evidence provides us with epistemic reasons for belief only if we can be blameworthy for violating purely epistemic norms—then a challenge for the normativist about evidence consists in spelling out the nature of the kind of blame that can be appropriate in response to purely epistemic failings, thereby allowing us to see why (2) is false.

There are, of course, other arguments against the normativity of evidence. In the concluding Sect. 5, for instance, I will mention the challenge of making sense of conflicts between purely epistemic norms and practical norms. A full defense of the normativity of evidence would require responding to such challenges as well. Furthermore, a full account of epistemic normativity would have to provide a

justification of our epistemic practices—of why we are justified in having a practice of holding each other answerable to purely epistemic norms, responding with reactive attitudes to their violation, and so on. Here I will not provide such a full account. My more modest aim in this paper is rather to make room for the plausibility of a purely epistemic kind of normativity by showing how it reveals itself in our practice of holding each other answerable to epistemic norms. I will return to these issues briefly in the concluding Sect. 5 to clarify what this paper achieves, and what it doesn't.

To get epistemic blame into focus, I will distinguish between *culpable* and *non-culpable* violations of epistemic norms, and I will discuss briefly how different accounts about responsibility for belief evaluate both cases (Sect. 4.1). The kind of blame that is still appropriate in the non-culpable cases will then give us an idea about how a normativist about evidence could understand epistemic blameworthiness (Sect. 4.2). This will help us to evaluate cases of *trivial* belief in which a purely epistemic norm was violated (Sect. 4.3). The relevant concept of blame will allow us to make sense of the idea that evidence still provides us with reasons to believe when no practical value hinges on what we believe. The proposed view about the normativity of evidence allows us to see that epistemic norms have normative significance even without being hostage to practical value.

4.1 Blameworthiness for Non-culpable Violations of Epistemic Norms

Consider the following distinction between *culpable* and *non-culpable* violations of epistemic norms. There are violations of epistemic norms that we could have reasonably avoided (culpable), and violations which we could not have reasonably avoided (non-culpable). It seems that many deniers of human-induced climate change have plenty of evidence available that should rationally convince them that they are wrong if they would consider their evidence more carefully. Their epistemic norm violation is thus culpable. Like someone who commits a moral wrong for egoistic motives, these people fail to make an effort of will they owe to others by failing to consider their evidence. The epistemic norm violation has severe consequences for other people, and the culpable deniers could have reasonably avoided the violation. Due to this, it can be appropriate to react to them with moral blame, like resentment or indignation: they do not merely violate an epistemic norm, but also a moral norm of belief-management.

Compare the culpable denier of human-induced climate change with the non-culpable denier. The latter person might be someone who grew up in a community where human-induced climate change is denied by everyone even in the face of sufficient evidence. That is, even if members of this group are presented with clear evidence, they remain unconvinced. Consider one of these non-culpable deniers who was trained from an early age on to believe against the evidence when it comes to the topic of climate change. Careful consideration of the evidence is not anything that comes up as a reasonable course of action to this denier, nor can we reasonably expect this of them. Knowing the social background of the denier, we can say that they lost some authority over their beliefs when it comes to human-induced

climate change.[16] This is why resenting them or being indignant does not seem to be appropriate.

Some recent accounts of responsibility for belief would argue that the non-culpable deniers are not blameworthy for their beliefs *in any sense*. They argue that our responsibility for belief is always derivative to our responsibility for actions and omissions prior to that belief by means of which we could have had some reasonable kind of influence or control over the belief (cf. Meylan, 2013, 2017; Peels, 2017)—say, actions of inquiry or investigation, or of actively considering one's evidence. Consequently, if there was no reasonable course of action for our denier that would have led them to adopt a different doxastic attitude—as we stipulated –, then these accounts imply that our denier is blameless.

Interestingly, given premise (1) of the argument from doxastic blameworthiness, these accounts are committed to denying the normativity of evidence—i.e., they are committed to ANE. For their account implies that a person who holds a blameworthy belief had reasons for actions or omissions by means of which they could have avoided their belief. If so, then a violation of purely epistemic norms does never, by itself, make the person blameworthy. Rather, according to these accounts, blameworthiness presupposes that there were reasons for actions and omissions by means of which the person could have managed their belief. Given (1), ANE follows from this claim.[17]

Other accounts of responsibility for belief would disagree, however. They would argue instead that even the non-culpable deniers might still be *answerable* for their beliefs (cf. Hieronymi, 2006, 2008, 2014; Smith, 2005, 2015). That is, they argue that it might still be intelligible to request of them to *justify* their beliefs by asking them for the evidence they take to bear on whether there is human-induced climate change. In contrast to brute headaches, the deniers' beliefs still reveal an aspect of their overall epistemic character due to being rationally evaluable. That is, as long as we assume that the non-culpable denier's beliefs are not wholly unresponsive to reasons, there seems to be a sense in which we could still be justified to react with negative attitudes towards them if they are incapable of providing a satisfying reply to our request for evidence. According to Hieronymi (2009), although the non-culpable denier's beliefs are not under their indirect *voluntary* control, they might still be conceived of as being under their *evaluative* control: their beliefs might still be active responses to their reason-giving environment (even though they are non-voluntary responses that are irrational), and can thus legitimately give rise to serious forms of criticism or blame.

This is not the place to decide which account of responsibility for beliefs—indirect control accounts or answerability-accounts—are right. Rather than deciding the dispute about what grounds our responsibility for belief, I will instead turn to my

[16] In the terminology employed by Fischer and Ravizza (1998) and McCormick (2015), we might say that they lost their *ownership* over the relevant belief-forming mechanisms.

[17] That these accounts are committed to a denial of the normativity of evidence might give rise to a strong argument against them. As far as I know, this implication has not been noticed before. I suppose that this is because the relationship between reasons and blameworthiness (or responsibility) is underexplored.

proposal about the nature of the blaming-responses that might still be appropriate towards the non-culpable deniers according to the answerability accounts. However, note that this will also pose a problem for indirect control accounts of doxastic responsibility. For if there *is* a kind of distinctively epistemic blame appropriate in the case of the non-culpable deniers, then they are epistemically blameworthy for their beliefs even though they could not have reasonably managed them by exercising indirect control. Most importantly for the present purposes, however, considering the nature of a purely epistemic kind of blame will give us a clue about how to understand the normativity of evidence.

4.2 Epistemic Blame, Relationships, Vices, and Trust

Hieronymi's (2004, 2019) and Smith's (2013) approaches to the nature of blame are in line with Scanlon's (1998, 2008) account. According to this family of accounts of the nature of blame, blaming someone need not mean that one feels emotions like resentment or indignation towards the person. Rather, we might blame someone merely by *modifying our relationship* towards them in a certain way. We might blame a person without feeling any hostility towards them, e.g., by just ceasing to be friends, or by no longer providing special support to the blamee, or by not taking pleasure in their successes, or by not valuing their opinions in the way we did before, or by developing a general sense of distrust towards them. Recently, especially Boult (2020, 2021) has applied these accounts to the epistemic domain. The following sketch of an account of epistemic blame draws on his ideas.[18]

Importantly, not all relationship modifications count as instances of blame. First, one might modify a relationship in a positive way, say, when one becomes so fond of someone that one wants to be closer friends with them; or when a parent finds out that their child committed a crime and in response to this cares even *more* about them (Smith, 2013, 137). Secondly, relationship modifications can happen without negative judgment about another person—as when people who live in different places just drift apart. Scanlonian approaches to blame thus owe us an account of what makes a *negative* relationship modification an instance of blame.[19]

[18] Cf. his cited works for a detailed defense. Other accounts of epistemic blame have been worked out by Kauppinen (2018), who also presents an account of epistemic criticism as a form of distrust, and by Brown (2020), who also shares the spirit of Boult and Kauppinen in that she regards epistemic blame as being neither a mere negative evaluation nor a kind of strong reactive emotion (like resentment). However, Brown builds on Sher's (2006, 2009) account of blame in order to spell out an epistemic kind of blame. Neither Kauppinen nor Brown put relationship modifications at center stage. Cf. Boult (2021) for a detailed discussion and critique of Brown's account.

[19] In response to Smith's case mentioned in the paragraph above, Boult (cf. 2021, 17) argues that genuine blaming reactions are those that are based on the judgment that the person is blameworthy. However, a problem with this proposal is that it does not tell us what our judgments about blameworthiness are based on. My proposal (that I explain in the next paragraph) avoids this problem: our judgments of blameworthiness are responses to the blamee's vice. Furthermore, it seems that the parent in Smith's case regards their child as blameworthy without blaming them. I think Smith's case can be met simply by restricting blaming responses to *negative* relationship modifications. Maybe Smith's (2013) proposal that blaming responses are expressions of one's protest is compatible with my proposal that they are based on

Negative relationship modifications count as instances of blame only if they are responses to the person's vice. I do not count as blaming my friend by judging and treating them as unreliable, and by modifying my expectations accordingly, if I am aware that their unreliability is due to factors that do not stem from their faulty character. Such factors might include their newborn child that makes them spontaneously cancel on me, or their depression that is the cause for their unreliability. Such factors do not give me a reason not to trust them, but merely a reason not to *rely* on them. By contrast, it might be legitimate to blame a friend if their unreliability indicates that they do not care about the friendship as much as one can reasonably expect of them as a friend. In this case, they are not fully honest about their attitude towards the friendship. Reducing one's trust in them, and thus modifying one's relationship with them negatively in response to their vice of dishonesty, can be legitimate.[20]

Negative relationship modifications in response to vices can plausibly count as blaming responses because they are only legitimate towards responsible beings. This is because we can only have the specific kind of relationship that is presupposed by these reactions with fully responsible beings. Neither computer nor children or animals can display vices that give rise to the negative reactions described above. Their misbehavior can only give rise to impoverished analogues of these reactions. For instance, I might "not trust" a dog in the sense that I suspect that they will bite me. The dog's behavior might be unreliable, but it won't give me a reason to blame the dog, since the dog's behavior does not manifest a vice (on the assumption that dogs cannot have full-blown vices like fully responsible beings). This indicates that negative relationship modifications in response to vices presuppose a subject's responsibility for their character and attitudes. At the same time, the appropriateness of these reactions does not presuppose that the subject could have managed their character or attitudes: these reactions merely presuppose an underlying vice, independently of its origin in voluntary conduct. Since these reactions presuppose responsibility, their potential "coolness" does not, *pace* Wallace (2011), count against them as genuine blaming-reactions (cf. Boult, 2020).[21]

Footnote 19 (continued)
epistemic vice—maybe epistemic blame expresses protest against epistemic vices. However, cf. Boult (2021) for some skepticism about the applicability of Smith's account to the epistemic domain.

[20] Cf. Smith's (2005, 242) case from George Eliot's *Scenes of Clerical Life* (1858), where Captain Wybrow fails to notice that Miss Assher never takes jelly, which, according to Smith, "suggests to Miss Assher that she does not yet occupy a distinctive place in his overall emotional and evaluative outlook" (2005, 243). Assher might legitimately modify her expectations towards Wybrow, and thus modify the relationship negatively, in response to Wybrow's vice of not caring about the relationship as much as he should. According to the proposed account, this would count as an instance of blame even if it does not involve emotions of resentment or indignation.

[21] An anonymous referee objected that negative relationship modifications might merely count as ways of *holding responsible*, but not of ways of *blaming*. This raises the question of how full-blown moral blaming responses like resentment and indignation relate to the "softer" Scanlonian negative reactive attitudes. For my purposes here, it is sufficient to note that if we hold a person responsible in the Scanlonian sense, this can provide the basis for a positive or a negative relationship modification in response to virtue or vice. For instance, we might want to be closer friends because of the person's virtues, or we might reduce our involvement with the person due to their vices. I take it to be of secondary interest whether we *call* these reactions "praise" and "blame". At the very least, they seem to be positive and negative ways of holding responsible. This is sufficient for these reactions to reveal the normative sig-

As David Owens puts it, after discussing the epistemic vice of gullibility: when I display a vice indicating a flaw in my character, then "I cannot be trusted to think and feel as I ought" (2000, 124). The normativity of these "oughts" is revealed, according to view I propose, by the fact that violating them impairs our relationship to others in specific ways so that it becomes appropriate to negatively modify one's relationship—e.g., by reducing one's presumption of epistemic trust. This impairment exists even if the person had no opportunity to manage their vice: as long as the epistemic vices are still genuine *vices* (rather than pathologies), non-culpable violations of epistemic norms that reveal a person's epistemic vice can impair our epistemic relationships, and thus give rise to suspension of epistemic trust.

If we allow for a broad concept of blame in terms of impaired relationships, then we might be able to make room for something like purely epistemic blameworthiness. In an initial attempt, we might state that if we are blameworthy *morally* as soon as our relationship to our *moral* community is impaired, then we are blameworthy *epistemically* as soon as our relationship to our *epistemic* community is impaired. This impairment might matter in specific ways for how we should relate to one another: whether we believe the other person, whether we provide them with information, and whether we engage with them in rational discourse.

One problem with this initial formulation is that one's moral or epistemic community can be epistemically or morally flawed, and thus one might end up impairing one's relationship with them by being morally or epistemically virtuous.[22] Boult's (2020) formulation of the position avoids this problem: one is blameworthy epistemically only if one falls short of the *normative ideal* of an epistemic relationship—or, in my preferred terminology, only if one displays an epistemic vice. The epistemically virtuous person does not fall short of this ideal even within an epistemically flawed community. Thus, members of the community won't have a reason to reduce their epistemic trust in the virtuous person. By appealing to the normative ideal of an epistemic relationship, we can explain why being dogmatic or gullible, engaging in wishful thinking, or being biased can make one epistemically blameworthy even in epistemic communities that socially reward such vices. For all these vices are, as Boult puts it, problematic ways of exercising one's epistemic agency that make one fall short of the normative ideal and thus warrant suspension of one's presumption of epistemic trust.[23]

Footnote 21 (continued)
nificance of a norm that we mark as violated by reacting in these ways. For defenses of such reactions as genuine *blaming* responses, cf. the recent works that argue that these reactions count as blame because they go hand in hand with, or consist in, a kind of motivation—a desire that the blamee had not "believed badly" (Brown, 2020), a protest against the blamee's action or attitude (Smith, 2013) or just generally the motivation to change one's relationship with them by modifying one's expectations and intentions (Boult 2021).

[22] I am grateful to an anonymous referee for making me aware of this problem.

[23] I take it that Boult does not use "epistemic agency" as referring to indirect voluntary control over beliefs. Plausibly, one can be dogmatic, gullible, a wishful thinker or biased even if this was not under one's indirect voluntary control. Rather, beliefs—including irrational ones—are often involuntary responses to one's environment (cf. Strawson, 2003). Boult's notion of epistemic agency is more plausibly understood in terms of Hieronymi's (2009) notion of evaluative control.

4.3 Blameworthiness for Non-culpable and for Trivial Violations of Epistemic Norms

Let us apply this sketch of an account of epistemic blame to our two relevant cases: *non-culpable* violations and *trivial* violations of epistemic norms. In both cases, it seems as if passionate forms of moral blame, like resentment or indignation, are no longer appropriate. Importantly, however, reducing our presumption of epistemic trust for failing to live up to the normative ideal of an epistemic relationship might still be appropriate. This will allow us to see that we already take epistemic norms to be normatively significant: we are already committedly involved in epistemic sociality.[24] Our actual practice of holding each other answerable to purely epistemic norms presupposes their normative significance insofar as we express this significance in our reactive attitudes towards norm violators.

The non-culpable deniers of climate change might still be manifesting an epistemic vice, a defect in character, which we might label "epistemic irrationality".[25] Note first that their epistemic irrationality is *attributable* to them in the sense that it is part of their overall outlook on the world, rather than just an occasional lapse which we could excuse. It is thus a genuine vice. Secondly, our non-culpable norm violators are still *answerable* for their beliefs insofar as it is intelligible to request their *evidence* for their beliefs. For we conceived of the case in such a way that their disbelief is still rationally evaluable rather than pathological: we assumed that their belief is irrational in the sense that they are aware, on some level, of the evidence against their belief, but they still fail to respond correctly to their overall evidence due to epistemic vices like dogmatism, gullibility, or a tendency for wishful thinking. Our epistemic blame directed at the non-culpable deniers is based on our judgment that they cannot give a satisfying answer to our request for evidence, even though the evidence is readily available to them. As a result, we have a reason to suspend our presupposition of epistemic trust towards them.[26]

What is the verdict, according to this account of epistemic blame, about our blameworthiness for trivial belief that is insufficiently supported by one's evidence? The normativist about evidence has two strategies available. Both strategies can be combined.

First, normativists could argue that even violations of epistemic norms in trivial matters might indicate a general flaw in the epistemic character of a person. As Boult puts it when evaluating trivial cases, "[s]o long as I modify my intentions and expectations towards them, in a way made fitting by the judgment (however implicit) that

[24] I owe this idea to Matthew Chrisman.

[25] On rationality as a virtue, cf. Wedgwood (2017). Although I do not wish to make any substantial claims about rationality here, note that the idea I present in this paragraph fits well with the argument that irrationality (which can be non-culpable) is criticizable and must therefore be connected to, or even spelled out in terms of, normative reasons (cf., e.g., Kauppinen, 2019, 3; Kiesewetter, 2017, chapter 2; Schmidt, forthcoming; Way, 2009, 1).

[26] Importantly, things would be rather different if the non-culpable deniers were not in possession of sufficient evidence for human-induced climate change. For instance, we might conceive of a community that managed to shut off any evidence from the outside world, and that worked on discrediting science within their community over decades. In this case, disbelieve in human-induced climate change might turn out to be rational for them, given their misleading epistemic perspective. In this case, we might not *rely* on their judgments. But we then won't count as blaming them—their unfortunate circumstances would excuse most individuals within the community. They would not count as epistemically vicious, because they are epistemically rational to a normal degree.

they've impaired the general epistemic relationship, then I count as epistemically blaming them" (Boult, 2020, 9). That is, if your friend tends to believe celebrity gossip that they read in a magazine they know to be unreliable, this might give you a (*pro tanto* or *prima facie*) reason to suspend epistemic trust in them. Presumably, this could mean that you should suspend your trust in some situations when it comes to matters of importance, because you now have some evidence that their epistemic character is flawed.

Secondly, the normativist can just grant that violations of epistemic norms do not *always* make it appropriate to suspend trust. For they need not argue that such violations *always* make one epistemically blameworthy. In order to disprove premise (2), it is enough to show that we are *sometimes* blameworthy in virtue of the fact that we violated a purely epistemic norm. More generally, violating a reason-providing norm need not amount to displaying a criticizable vice, and thus need not amount to blameworthiness. Compare the idea that someone's morally wrong action is not necessarily blameworthy. We all act wrong from time to time, and we all violate epistemic norms from time to time. We can usually *excuse* each other for occasional lapses and do not regard these lapses as having any significant consequences for our interpersonal relationships. Yet moral wrongs and violations of epistemic norms are lapses nevertheless—i.e., they are violations of norms that provide us with reasons for compliance.

Seeing that reducing epistemic trust is an appropriate negative response to an epistemic vice and that it marks the impairment of an epistemic relationship provides us with a plausible starting point for understanding the significance of epistemic normativity. It allows us to meet the challenge for the normativity of evidence presented in this paper by rejecting premise (2). This challenge claims that the absence of a distinctively epistemic kind of blame rules out the normativity of evidence. I have proposed that we can meet this challenge by appealing to recent accounts of doxastic responsibility as answerability and to recent accounts of epistemic blame. The former accounts show us that non-culpable beliefs might still be blameworthy. The latter accounts provide us with an idea about what this blameworthiness could consist in, and how it could sometimes extend also to cases of trivial belief. By building on Boult's account, I have suggested that epistemic blame consists in marking impaired epistemic relationships by reducing epistemic trust in response to a person's epistemic vice.

5 Conclusion and Outlook

The dispute about the normativity of evidence is a currently lively discussion within epistemology.[27] There is no need right now to settle the dispute once and for all. This paper has contributed two ideas towards bringing the debate forward:

[27] Cf. also the recent discussion about the normativity of epistemic reasons. Cf. Kiesewetter (forthcoming) for an overview over this debate and defense of the normativity of epistemic reasons. According to Kiesewetter, some instrumentalists or pragmatists are best read as making room for a notion of a non-normative epistemic reason, rather than as denying the existence of epistemic reasons altogether. My terminology here does not allow us to distinguish between these two positions, because I used 'epistemic reason' exclusively in a normative sense.

(a) That a central challenge in defending the normativity of evidence consists in spelling out a notion of a distinctively *epistemic* kind of blame.
(b) That appealing to epistemic blame as marking an impaired epistemic relationship by suspending epistemic trust in response to epistemic vices is a promising way for defending the distinctive normativity of evidence in the sense required by this challenge.

Together, both claims shift the dialectical burden towards ANE. Proponents of ANE argue that purely epistemic norms do not provide us with reasons to believe. But if epistemic blame marks the violation of purely epistemic norms with reason-providing force, then proponents of ANE must say that there is no such thing as a distinctively epistemic kind of blame. However, recent approaches on doxastic responsibility as answerability (Hieronymi, Smith), as well as recent works that spell out the nature of a distinctively epistemic kind blame or criticism (Boult, Brown, Kauppinen), call this into doubt. As a result, proponents of ANE need to engage with these theories: they have to show why the appropriateness of the blaming-reactions that these theories spell out does not imply that a norm with reason-providing force was violated; or else argue that these reactions are not appropriate in response to violations of purely epistemic norms. However, I have suggested in Sect. 4 that violations of epistemic norms in non-culpable and trivial cases can well deserve suspension of epistemic trust if they are manifestations of epistemic vice. The presented analysis of these cases thus calls into doubt ANE by revealing a purely epistemic kind of blame that might be appropriate in these cases—a blame that reflects the normative significance we attach to purely epistemic norms.

However, this does not yet provide us with a full account of epistemic normativity. I will now briefly explain what I think such an account requires, at a minimum. This will reveal the restrictions of the present inquiry. At the same time, it illustrates how the approach presented in this paper might be fruitfully developed to a fuller account of epistemic reasons.

The first requirement for a full account of epistemic normativity is that it must allow us to meet other challenges for the normativity of evidence. For instance, how do normativists about evidence deal with cases in which complying with an epistemic norm causes practical disvalue—for instance, cases in which others will suffer harm unless I make myself violate an epistemic norm? The proposed view about epistemic blame might help us to make sense of the traditional verdict that I *ought epistemically* to comply with the epistemic norm even though I *ought practically* to bring myself not to comply with it. Proponents of ANE will argue that the first "ought" cannot by normatively authoritative in any interesting sense. The proposed view, by contrast, allows us to say that the first epistemic "ought" has still a kind of normative authority insofar as the normative significance of this "ought" is expressed in the fact that members of one's epistemic community might be justified in modifying their trusting attitude towards me if I do not comply with it. That is, even if I bring myself to violate an epistemic norm for good practical reasons, I might end up not being trustworthy epistemically due to the resulting ill-based belief. This will hold at least in cases where my resulting ill-based belief reflects an

epistemic vice. However, the discussion about such cases is currently very alive, and this paper has no ambitions meeting this and further possible challenges.[28]

Secondly, this paper did not provide an account of the *source* of epistemic normativity. My appeal to our actual practice reveals that we treat each other as answerable to epistemic norms, and thus that we attach normative significance to these norms. But this does not justify our commitment to the overall epistemic practice. Indeed, the view proposed here is even compatible with a *pragmatic* foundation of purely epistemic norms: maybe we are justified to engage in our epistemic practice because it is practically valuable to be subject to epistemic norms (cf. Owens, 2017). Combined with such a pragmatic justification of our overall practice, normativists about evidence could maintain that *within* this practice, all reasons for belief are provided by our evidence, and that pragmatic considerations are only relevant if we wish to externally justify our adherence to this purely evidential kind of normativity.[29] This might be an important element in a complete error-theory about pragmatist-instrumentalist intuitions concerning reasons for belief.

Acknowledgements Thanks to two anonymous referees for very helpful comments.

Funding Open Access funding provided by Universität Zürich.

References

Berker, S. (2018). A combinatorial argument against practical reasons for belief. *Analytic Philosophy, 59*, 427–470.

Bondy, P. (2018). *Epistemic rationality and epistemic normativity*. Routledge.

[28] Heil (1992) gives a helpful impression of the two opposing intuitions in such cases: (a) there are two incommensurable or incomparable senses of "ought" *versus* (b) only the practical ought matters. For recent defenses and elaborations along the lines of (b), cf. Howard (2020), Maguire and Woods (2020), Meylan (2020), Reisner (2008), Rinard (2017), Steglich-Petersen and Skipper (2019). For verdicts along the lines of (a), cf., next to Heil (1992), Berker (2018) Feldman (2000, 680–681), Kelly (2003, 619), Pojman (1993).

[29] Cf. this view with Dennett's parallel view about the justification of punishment: while a practice of punishment can be externally justified, according to Dennett, only by consequential considerations (deterrence, re-socialization, etc.), our particular judgments about who is to be punished are justified by desert-based considerations (who deserves to be punished, whom it is unfair to punish, etc.) that are internal to this practice (cf. Dennett & Caruso, 2021, 119–127). Analogously, our beliefs might be justifiable only by purely evidential considerations while our overall practice must be justified on pragmatic grounds.

Boult, C. (2020). There is a distinctively epistemic kind of blame. *Philosophy and Phenomenological Research*. https://doi.org/10.1111/phpr.12726
Boult, C. (2021). The significance of epistemic blame. *Erkenntnis*, 1–22. https://doi.org/10.1007/s10670-021-00382-0
Brown, J. (2020). What is epistemic blame? *Noûs, 54*, 389–407.
Cowie, C. (2014). In defence of instrumentalism about epistemic normativity. *Synthese, 191*, 4003–4017.
Dennett, D., & Caruso, G. D. (2021). *Just deserts*. Polity Press.
Feldman, R. (2000). The ethics of belief. *Philosophy and Phenomenological Research, 60*, 667–695.
Feldman, R. (2007). Reasonable religious disagreements. In L. Antony (Ed.), *Philosophers without god: Meditations on atheism and the secular life* (pp. 194–214). OUP.
Fischer, J., & Ravizza, M. (1998). *Responsibility and control. A theory of moral responsibility*. CUP.
Harman, G. (1986). *Change in view: Principles of reasoning*. MIT Press.
Heil, J. (1992). Believing reasonably. *Noûs, 26*, 47–62.
Hieronymi, P. (2004). The force and fairness of blame. *Philosophical Perspectives, 18*, 115–148.
Hieronymi, P. (2006). Controlling attitudes. *Pacific Philosophical Quarterly, 87*, 45–74.
Hieronymi, P. (2008). Responsibility for believing. *Synthese, 161*, 357–373.
Hieronymi, P. (2009). Two kinds of agency. In L. O'Brian & M. Soteriou (Eds.), *Mental actions* (pp. 138–168). OUP.
Hieronymi, P. (2014). Reflection and responsibility. *Philosophy and Public Affairs, 42*, 3–41.
Hieronymi, P. (2019). I'll bet you think this blame is about you. In D. J. Coates & N. Tognazzini (Eds.), *Oxford studies in agency and responsibility V: Themes from the philosophy of Gary Watson* (pp. 60–87). OUP.
Hofmann, F. (2020). Is evidence normative? *Philosophia*. https://doi.org/10.1007/s11406-020-00241-2
Howard, C. (2020). Weighing epistemic and practical reasons for belief. *Philosophical Studies, 177*, 2227–2243.
Kauppinen, A. (2018). Epistemic norms and epistemic accountability. *Philosopher's Imprint, 18*, 1–16.
Kauppinen, A. (2019). Rationality as the rule of reason. *Noûs*. https://doi.org/10.1111/nous.12311
Kelly, T. (2003). Epistemic rationality as instrumental rationality: A critique. *Philosophy and Phenomenological Research, 66*, 612–640.
Kiesewetter, B. (2017). *The normativity of rationality*. OUP.
Kiesewetter, B. (forthcoming). Are epistemic reasons normative? *Noûs*.
Lord, E. (2018). *The importance of being rational*. OUP.
Maguire, B., & Woods, J. (2020). The game of belief. *The Philosophical Review, 129*, 211–249.
Mantel, S. (2019). Do epistemic reasons bear on the ought simpliciter? *Philosophical Issues, 29*, 214–227.
McCormick, M. S. (2015). *Believing against the evidence*. Routledge.
McCormick, M. S. (2020). Believing badly: Doxastic duties are not epistemic duties. In K. McKain & S. Stapleford (Eds.), *Epistemic duties: New arguments, new angles* (pp. 29–43). Routledge.
Meylan, A. (2013). *Foundations of an ethics of belief*. Ontos.
Meylan, A. (2017). The consequential conception of doxastic responsibility. *Theoria, 83*, 4–28.
Meylan, A. (2020). Doxastic divergence and the problem of comparability. Pragmatism defended further. *Philosophy and Phenomenological Research*. https://doi.org/10.1111/phpr.12709
Owens, D. J. (2000). *Reason without freedom*. Routledge.
Owens, D. J. (2017). *Value and epistemic normativity, in his normativity and control* (pp. 87–109). OUP.
Papineau, D. (2013). There are no norms of belief. In T. Chan (Ed.), *The aim of belief* (pp. 64–79). OUP.
Peels, R. (2017). *Responsible belief*. OUP.
Pojman, L. (1993). Believing, willing, and the ethics of belief. In L. Pojman (Ed.), *The theory of knowledge* (pp. 525–543). Wadsworth.
Reisner, A. (2008). Weighing pragmatic and evidential reasons for belief. *Philosophical Studies, 138*, 17–27.
Rinard, S. (2015). Against the new evidentialists. *Philosophical Issues, 25*, 208–223.
Rinard, S. (2017). No exception for belief. *Philosophy and Phenomenological Research, 94*, 121–143.
Scanlon, T. M. (1998). *What we owe to each other*. HUP.
Scanlon, T. M. (2008). *Moral dimensions: Permissibility, meaning, blame*. Belknap Press of Harvard University Press.
Schmidt, S. (forthcoming). Rationality and responsibility. *Australasian Philosophical Review, 5*.
Sher, G. (2006). *In praise of blame*. OUP.
Sher, G. (2009). *Who knew? Responsibility without awareness*. OUP.

Smith, A. M. (2005). Responsibility for attitudes: Activity and passivity in mental life. *Ethics, 115*, 236–271.
Smith, A. M. (2013). Moral blame as moral protest. In D. J. Coates & N. Tognazzini (Eds.), *Blame: Its nature and norms* (pp. 27–48). OUP.
Smith, A. M. (2015). Responsibility as answerability. *Inquiry, 58*, 99–126.
Steglich-Petersen, A. (2011). How to be a teleologist about epistemic reasons. In A. Steglich-Petersen & A. Reisner (Eds.), *Reasons for belief* (pp. 13–33). CUP.
Steglich-Petersen, A., et al. (2018). Epistemic instrumentalism, permission, and reasons for belief. In C. McHugh (Ed.), *Normativity: Epistemic and practical* (pp. 260–280). OUP.
Steglich-Petersen, A., & Skipper, M. (2019). An instrumentalist account of how to weigh epistemic and practical reasons for belief. *Mind, 129*, 1071–1094.
Steglich-Petersen, A., & Skipper, M. (2020). Instrumental reasons for belief: Elliptical talk and elusive properties. In S. Schmidt & G. Ernst (Eds.), *The ethics of belief and beyond* (pp. 109–125). Routledge.
Strawson, G. (2003). Mental ballistics or the involuntariness of spontaneity. *Proceedings of the Aristotelian Society, 103*, 227–256.
Stroud, S. (2006). Epistemic partiality in friendship. *Ethics, 116*, 498–524.
Wallace, R. J. (2011). Dispassionate opprobrium: On blame and the reactive sentiments. In R. J. Wallace, R. Kumar, & S. Freeman (Eds.), *Reason and recognition: Essays on the philosophy of T. M. Scanlon.* OUP.
Way, J. (2009). Two accounts of the normativity of rationality. *Journal of Ethics and Social Philosophy, 4*, 1–9.
Wedgwood, R. (2017). *The value of rationality*. OUP.
White, R. (2005). Epistemic permissiveness. *Philosophical Perspectives, 19*, 445–459.
Whiting, D. (2010). Should I believe the truth? *Dialectica, 64*, 213–224.
Whiting, D. (2013). Truth: The aim and norm of belief. *Teorema, 32*, 121–136.

Journal of Environmental Law, 2021, 33, 669–696
doi: 10.1093/jel/eqab018
Original article

OXFORD

Improving Energy Efficiency: The Significance of Normativity

Elizabeth A. Kirk * and Laurel Besco**

ABSTRACT

The failure of the global community to effectively address many large-scale environmental challenges calls into question the existing regulatory approaches. A large number of these challenges are diffuse issues which have, over the years been targeted by significant and sizable regulatory frameworks and yet the challenges persist—energy efficiency is one such issue and is the focus of this article. Increasing monitoring or enforcement to achieve improvements in regulatory compliance is too expensive in the context of diffuse problems due to the scale and costs such activities would entail. We suggest a focus on the fit between regulatory frameworks and norm creation may identify more fruitful routes to regulatory reform. Drawing on the 'interactional account of law' as a framework, this research uses new empirical data from a survey and a set of interviews to investigate the failure of energy efficiency regulatory frameworks at achieving energy efficient norms of behaviour in industry. We look at Canada and the UK as our case studies and our emphasis is on industry actors as they represent a significant and yet understudied area of society. We find that though existing regulatory structures seem adequate to generate general shared understandings around obligations to engage in energy efficiency actions, more specific shared practice around actually engaging in these actions remains elusive, resulting in a failure to engender norms of behaviour. These failures, we suggest, link directly to an inadequate fit between the regulatory tools and Fuller's criteria for the internal morality of law.

KEYWORDS: corporate action, normativity, interactional account of law, internal morality of law, government intervention, energy efficiency

1. INTRODUCTION

Many pressing global sustainability issues are diffuse in nature. Unfortunately, many of these same issues are also associated with regulatory failure, with targets for improvement often not being met. While increased monitoring and enforcement of existing regulatory tools is one solution, the diffuse nature of these issues means that most actions take place behind company doors, and so the costs of enforcement are

* Director, Lincoln Centre for Ecological Justice, University of Lincoln, Brayford Pool, Lincoln, LN6 7TS, UK. (ekirk@lincoln.ac.uk).

** Department of Geography, Geomatics and Environment and Institute for Management and Innovation, University of Toronto Mississauga, Mississauga, ON, L5L 1C6, Canada. (laurel.besco@utoronto.ca).

likely to be excessive.[1] One route to improving regulation is to adopt tweaks to existing measures to improve their efficacy.[2] We propose that a more fruitful approach may be to deepen our understanding of why the regulated do not always appear to act as though they are rational actors, despite much of our current regulation being rooted in this conception of corporate actors and industry.[3] We draw on Brunnée and Toope's interactional account of law[4] for this purpose. Although developed in the context of international law, we demonstrate that this theory also illuminates the causes of the gap between regulatory intention and concrete action at the national level and does so by distinguishing between formal regulation/regulatory instruments and norms. (In the context of our research, we understand a norm as representing 'desirable behaviors [*sic*] for a population of a natural or artificial community . . . generally understood as rules indicating actions that are expected to be pursued that are either obligatory, prohibitive, or permissive based on a specific set of facts.')[5] When examining diffuse problems, such as energy efficiency, the interactional account of law allows us to see more clearly that the causes of regulatory failure may be rooted in the resource demands (be that financial, human, time or other resources) that regulations place on companies. Thus, corporations may indeed wish to behave as rational actors, but resource constraints prevent them from doing so. These findings we suggest may also illuminate a route to improving regulatory effectiveness.

To draw out the gaps between norms, regulation and action by industry actors in relation to energy efficiency, we have engaged in empirical research focused on the perspective of industry actors in Canada and the UK. In doing so, we address a gap in the literature in relation to the role of governments in the creation of normativity. While writers such as Green[6] and Sunstein[7] suggest that governments have a role to play in norm formation, they do so in the context of individuals. Industry actors themselves are less often the focus of this type of research[8] and yet it is essential to truly understand how obligations are felt by industry and what action those obligations do or do not spur. We use an empirical study of three industries—building and construction (BC), hospitality (H) and energy utilities (E) —across Canada and the

1 Neil Gunningham, 'Environment Law, Regulation and Governance: Shifting Architectures' (2009) 21 Journal of Environmental Law 179.

2 Andrew Newman, 'The Green Corporate Citizen-Renovating the Corporation to Institutionalise Environmental Sustainability' (2013) 15 Asia Pacific Journal of Environmental Law 125.

3 Ronald H Coase, 'The Problem of Social Cost' (1960) 3 Journal of Law and Economics 1; Ian Ayres and John Braithwaite, *Responsive Regulation: Transcending the Deregulation Debate* (Oxford University Press 1992).

4 Jutta Brunnée and Stephen J Toope, *Legitimacy and Legality in International Law: An Interactional Account* (Cambridge University Press 2010), vol 67.

5 Moamin A Mahmoud and others, 'A Review of Norms and Normative Multiagent Systems' [2014] The Scientific World Journal 2014.

6 Andrew Green, 'Norms, Institutions, and the Environment' (2007) 57 University of Toronto Law Journal 105.

7 Cass R Sunstein, 'Endogenous Preferences, Environmental Law' (1993) 22 Journal of Legal Studies 217; Cass R Sunstein, 'Social Norms and Social Roles' (1996) 96 Columbia Law Review 903.

8 Two notable exceptions to this are Marie Byskov Lindberg, Jochen Markard and Allan Dahl Andersen, 'Policies, Actors and Sustainability Transition Pathways: A Study of the EU's Energy Policy Mix' (2019) 48 Research policy 103668; Laurel Besco and Elizabeth A Kirk, 'Industry Perceptions of Government Interventions: Generating an Energy Efficiency Norm' (2020) 23 Journal of Environmental Policy & Planning 130.

UK to develop an understanding of the opinions of those bound by obligations and norms. Though empirical research in environmental law scholarship is growing,[9] it remains the exception and not the rule[10] and therefore this article also addresses this dearth of experience. Further, there is some evidence that where data are employed in environmental law research, it often relies on existing data rather than generating new data.[11] Therefore, in addition to the contribution to understanding of how regulation operates in theory and practice, this article contributes to building up scholarship both in the general empirical environmental law sense, but also in terms of research that generates its own data. As Fischman and Barbash-Riley note, '[a]n empirical agenda could facilitate reforms to improve environmental law's effectiveness'[12] and this is exactly our goal in this research—to understand and improve the effectiveness and compliance with energy efficiency law.

We begin, therefore, by explaining the suitability of Brunnée and Toope's interactional account of law to illuminate the challenges the regulated face in addressing diffuse problem scenarios such as addressing energy efficiency. We then discuss our research methods before moving on to provide a brief overview of energy efficiency laws and policies in Canada and the UK to give a better understanding of existing regulatory frameworks in these two countries. We then progress to an assessment of industry views on obligations for energy efficiency actions, drawn from the results of our empirical survey. Finally, we dive deeper into our discussion of industry views on existing energy efficiency law and policy, through the lens of Fuller's eight criteria. We do so by drawing on findings from interviews with industry actors.

2. THE INTERACTIONAL ACCOUNT OF LAW

There is a substantial body of literature on motivating corporate interests to engage in actions related to environment and sustainability[13] and so one may ask why it is necessary to turn to a theory from international law to improve understanding of failures in domestic regulation. One issue is that there is 'little

9 See Robert L Fischman and Lydia Barbarsh-Riley, 'Empirical Environmental Scholarship' (2017) 44 Ecology Law Quarterly 767. For examples of such scholarship, see also Joel A Mintz, 'Neither the Best of Times nor the Worst of Times: EPA Enforcement during the Clinton Administration' (2005) 35 Environmental Law Reporter: News & Analysis 10390; Michael W Toffel and Jodi L Short, 'Coming Clean and Cleaning up: Does Voluntary Self-Reporting Indicate Effective Self-Policing?' (2011) 54 The Journal of Law and Economics 609; Amy Lawton, 'Nudging the Powerful: Reflecting on How to Make Organisations Comply with Environmental Regulation' (2020) 32 Journal of Environmental Law 25.

10 See Lawton, ibid 27.

11 Research conducted on empirical studies in US Law Journals highlights this point. See Fischman and Barbash-Riley (n 9) 782.

12 ibid 768.

13 See Forest Reinhardt, 'Market Failure and the Environmental Policies of Firms: Economic Rationales for "beyond Compliance" Behavior' (1999) 3 Journal of Industrial Ecology 9; Craig Deegan, Michaela Rankin and John Tobin, 'An Examination of the Corporate Social and Environmental Disclosures of BHP from 1983-1997: A Test of Legitimacy Theory' (2002) 15 Accounting, Auditing & Accountability Journal; John L Campbell, 'Why Would Corporations Behave in Socially Responsible Ways? An Institutional Theory of Corporate Social Responsibility' (2007) 32 Academy of management Review 946; Caroline D Ditlev-Simonsen and Atle Midttun, 'What Motivates Managers to Pursue Corporate Responsibility? A Survey among Key Stakeholders' (2011) 18 Corporate Social Responsibility and Environmental Management 25; Hamish Van der Ven, 'Socializing the C-Suite: Why Some Big-Box Retailers Are "Greener" than Others' (2014) 16 Business and Politics 31; Lawton (n 9).

consensus about what secures high levels of business compliance with environmental or other regulatory laws.'[14] Despite this, as noted in Section 1, much, if not all of the regulations addressing corporate actions in relation to diffuse problems, such as how to improve energy efficiency, are rooted in the idea of corporate actors as rational actors. This type of approach has been counted upon to engender greater reliance on renewable energy.[15] Tools such as corporate reporting are used to encourage behaviour change in companies.[16] The Companies Act 2006 (Strategic Report and Directors' Report) Regulations 2013,[17] eg requires reporting on corporate greenhouse gas emissions which may prompt companies to adopt energy efficiency measures to reduce their emissions. However, as Larrinaga *et al* demonstrate, such disclosure requirements do not always lead to a change in actual corporate behaviour[18] and instead may be used to enhance public relations rather than to engender change within the organisation.[19] In this context, we suggest that Brunée and Toope's comment that 'without a rich understanding of how international law influences the behaviour of key actors, one cannot design effective political and legal strategies to accomplish shared or even individual goals'[20] is equally apposite in the context of domestic law focused on diffuse problems such as energy efficiency and the behaviour of industry actors. In other words, if we do not understand how laws and regulations influence the behaviour of, eg industry actors, we will struggle to develop laws and regulations that succeed in delivering the policy goals they are designed to implement.

Brunnée and Toope's interactional account of law enables one to distinguish more clearly between what formal regulatory instruments require, and the perceptions of the regulated regarding whether or not they are subject to an obligation to act. In Brunnée and Toope's words '[i]nteractional law helps us to understand that the formal indicator of a rule . . . is not necessarily co-extensive with the legality and practice that generates obligation.'[21] Put another way, simply because a statutory instrument, or a piece of legislation comes into force, does not mean that those it is directed at will change their behaviour.

Brunnée and Toope point to three elements being necessary for a binding obligation, or norm, to emerge. These are a shared understanding of the norm, consistency in practice and that any law embodying the norm must comply with Fuller's eight criteria for the 'internal morality' of law[22]: law must be promulgated, law must be

14 Peter Kellett, 'Securing High Levels of Business Compliance with Environmental Laws: What Works and What to Avoid' (2020) 32 Journal of Environmental Law 179.

15 Nadia B Ahmad, 'Responsive Regulation and Resiliency: The Renewable Fuel Standard and Advanced Biofuels' (2018) 36 Virginia Environmental Law Journal 40.

16 Barnali Choudhury and Martin Petrin, 'Corporate Governance That "Works for Everyone": Promoting Public Policies through Corporate Governance Mechanisms' (2018) 18 Journal of Corporate Law Studies 381.

17 The Companies Act 2006 (Strategic Report and Directors' Report) Regulations 2013, SI 2013/1970.

18 Carlos Larrinaga-González and others, 'The Role of Environmental Accounting in Organizational Change-An Exploration of Spanish Companies' (2001) 14 Accounting, Auditing & Accountability Journal 213.

19 Olaojo Aiyegbayo and Charlotte Villiers, 'The Enhanced Business Review' [2011] Journal of Business Law 699.

20 Brunnée and Toope (n 4) 6.

21 ibid 8.

general, law must not be retrospective, law must be clear and intelligible, law must be consistent, law must be free of contradictions, law must be possible to obey and laws must be administered congruently with their intent. Meeting these eight criteria is necessary for the law to be viewed as legitimate and thus for it to exert a compliance pull[23] and for practice and understanding to converge around the law, or regulation in question.

The need for shared understanding and consistency in practice as expressed in the interactional account of law builds on Fuller's idea that law depends on social practice for its formation and maintenance.[24] In other words, the passing of a law or regulation is not the end point in the legal process, but rather a part of a process which will either lead to the adoption and continuation of the norms expressed in the law, or lead to a change or end to those norms. In other words, those subject to laws must also have some involvement in their creation, or continued application, for the laws to be seen to create binding obligations across time. In Brunnée and Toope's theory that social practice is evidenced through shared understanding and consistent practice. The starting point is that there must be a shared understanding of the need for the law or obligation. The new law may stretch existing understanding regarding the obligation it embodies and aim to create new normative standards, but for it to be effective it must meet Fuller's eight criteria for the 'internal morality' of law and be followed by a coalescing of consistent practice around the new norms expressed in the law. In the case of energy efficiency, the focus of this article, this would mean that the regulatory tools in place must both create shared understandings of obligations and lead to shared practice of energy efficiency actions in order to engender norms of behaviour that would result in successful achievement of the regulatory framework's goals. As we demonstrate in this article, adherence to Fuller's eight criteria is crucial to the success of this process.

It may seem that Fuller's criteria hardly need restating. It is obvious, eg, that if law is not promulgated and not accessible to its subjects, it will be down to chance whether practice coincides with it. Where practice does not (on the whole) coincide with the law, citizens and others will object to enforcement action in respect of those laws. It is also fairly obvious that such objections and practices might lead to an erosion of trust in the law. Nevertheless, we posit that adhering to Fuller's eight criteria and considering their relationship with shared understanding and constant practice is key to regulatory success. Despite this, there are gaps in understanding. There is, eg little if any detail on the understanding of what those subject to, eg energy efficiency regulation, know about new regulations when they are introduced and so the question of the promulgation of law remains live. The same can be said of Fuller's other criteria. What does it mean, eg to say that laws must be general when we discuss the concept of energy efficiency? Must they be seen to apply to all parts of industry and all regions in which a particular industry operates? It might seem clear that laws must not be

22 Lon L Fuller, 'Positivism and Fidelity to Law: A Reply to Professor Hart' (1958) 71 Harvard Law Review 630; Lon L Fuller, *The Morality of Law* (rev edn, Yale University Press 1969).

23 Brunée and Toope (n 4) 27.

24 Fuller, *The Morality of Law* (n 22).

retrospective, and it is all well and good for those with legal training to read regulations to determine their clarity, but what can industry tell us about which laws are 'clear and intelligible' to them, or in their view free of contradictions? Does industry view energy efficiency laws as being applied consistently across time? And do their views on this question tell us anything about the time period to be considered when making, applying and reviewing laws or regulations for them to be viewed by the regulated as consistent across time? One might assume that energy efficiency laws would be possible to obey, but again, what might industry tell us about the challenges they face in complying with energy efficiency obligations? Might other factors limit their ability to comply, rendering the obligations impossible to obey? Lastly, what are we to understand by the idea that laws are to be administered congruently with their intent, when addressing diffuse issues such as energy efficiency where most actions take place behind company doors. Given the costs and challenges of enforcing laws that tackle diffuse problems,[25] might there be a risk that industry does not see officials acting congruently with the law as enforcement is not obvious, and might there, therefore, be a risk that the law will lack legitimacy?

As Brunnée and Toope note 'the interactional understanding of law also demonstrates that despite diversity, thin initial commitments to legality are possible and shared understandings may deepen through mutual engagement in communities of practice.'[26] In other words, new regulation, even if it meets Fuller's eight criteria, may on its own not be sufficient to change practice, but may lead to binding norms *if communities of practice can be engendered alongside* shared understanding. These three elements combined—Fuller's internal morality of law, shared understanding and constant/consistent shared practice—are what lead to norms that are felt to be binding by those they are directed at and so lead to widespread changes in behaviour.

3. RESEARCH METHODS

Brunnée and Toope's interactional account of law provides a framework for teasing out the existence, or lack of normativity in the context of energy efficiency measures. In so doing, it may help explain the differences between energy efficiency levels seen in Canada and the UK despite the relative similarity in their regulatory measures and their socio-legal context. For example, the 2018 International Energy Efficiency Scorecard[27] indicates that the UK has had greater relative success in delivering improvements in energy efficiency than Canada scoring a 5 out of a possible 6, compared with the Canada's 3 for the period 2010–15.[28] While the scorecard appears at first to demonstrate that the UK has stronger energy efficiency policies and regulations,[29] closer inspection indicates that the differences in policy and regulation

25 Gunningham (n 1).

26 Brunnée and Toope (n 4) 353.

27 Fernando Castro-Alvarez and others, 'The 2018 International Energy Efficiency Scorecard' [2018] Report 1801, American Council for an Energy-Efficient Economy, Washington, DC.

28 ibid 22–23.

29 The UK scores a maximum of 3 points for energy efficiency goals, while Canada scores only a 1 out of possible 3 (ibid 26), but this difference may be explained in the way the goals are expressed.

between the two countries[30] are not as significant as their differences in overall energy efficiency.[31] This suggests, therefore, that the regulated in Canada are not responding to the policies and regulation they are subject to as effectively as those in the UK are. We used two approaches to disentangle these issues. An online survey (detailed below) was used to illuminate generally how energy efficiency was viewed and to establish whether there were clear distinctions between the two countries, or a single 'energy efficiency norm' based on shared understanding and practice across the two countries. To draw out deeper understandings of why current regulatory frameworks are, or are not succeeding in engendering normativity for energy efficiency actions, we used semi-structured interviews, also detailed below.[32]

As mentioned, Canada and the UK were chosen as the countries for study as they have similar socio-legal backgrounds, something which can help limit outside factors skewing results. Here our focus is on industry and specifically on three distinct industrial sectors—BC, energy generation/distribution (E) and H.[33] The emphasis on industry perspectives is itself somewhat unique, as there are limited studies that investigate this perspective.[34] Yet, without truly understanding how those on the frontlines of industrial action on energy efficiency feel about their obligations generally and with regards to specific actions, there is a significant gap in our understanding which may lead to ineffective government interventions. These three sectors were chosen as the focus of this study as they have been observed to differ in terms of their awareness and engagement with energy efficiency. For example, within the building sector, a large proportion of respondents to a national survey felt energy efficiency was important, but only a minority felt able to actively grow energy efficient construction unless changes in the building code are made.[35] The hotel sector has been active in pursing

30 For example, in building construction and renovation, the UK scores more highly for its regulations than Canada (13 out of 18 to 10 out of 18 (ibid 35), but the appliance ratings scores for the two countries are the same (4 out of 5) (ibid 51) and in other areas, such as thermal power plants, and the market for business providing energy efficiency services Canada is ahead of the UK (ibid 28–29).

31 Canada uses 2.47 kwh/U of GDP compared to that of the UK's 0.86 kwh (See Hannah Ritchie and Max Roser, 'Canada: Energy Country Profile' (*Our World in Data* 2020) <https://ourworldindata.org/energy/country/canada?country=CAN~GBR> accessed 26 March 2021) nor in improvements in energy efficiency. (Canada's energy intensity rates for its economy have reduced by 28% between 1995 and 2015 and the UK's by 47% over the same period (See Hannah Ritchie and others, 'Measuring Progress towards the Sustainable Development Goals' <https://sdg-tracker.org/energy#targets> accessed 26 March 2021.)

32 Prior to beginning data collection, ethics approval from the University of Toronto Research Ethics Board (Protocol Number 33665) was given. All participants consented to being part of the study—phone interviewees provided verbal consent that was recorded by the interviewer and consent for in-person interviews was collected through a signed consent form. Survey consent was gained by participants agreeing to the consent and information letter which formed the first page of the online survey.

33 The classification of participants into the three sectors was self-reported for interview respondents, but for surveys, North American Industry Classification System classification was used in Canada (Accommodation and food services; Construction; Utilities (including power generation, energy distribution, etc.) and Nomenclature of Economic Activities classification was used in the UK (Accommodation and food service activities, Construction, Electricity, Gas, Steam and air conditioning).

34 Lindberg and others (n 8); Besco and Kirk (n 8).

35 Jaimin Guillot, 'CEAA 2014 Survey: Canadian Business Attitudes on Energy Efficiency' (5 June 2014) <https://www.slideserve.com/jaimin/ceea-2014-survey-canadian-business-attitudes-on-energy-efficiency> accessed 7 June 2021.

environmental goals (eg, 17 well-known hotels in New York City have made a 'green pledge' focusing on reducing greenhouse gas emissions,[36] but whether they have also engaged in energy efficiency efforts is unclear). Electricity generators, as providers of energy, have already demonstrated high awareness of energy efficiency. Therefore, we might expect to see varying responses from actors in these different sectors.

3.1 Survey Methods

We used a survey, distributed through an online survey panel organised by a survey company[37] to gain insights into energy efficiency actions. The respondents are recruited by the survey company to a general panel, provide basic information of their characteristics and then are periodically offered surveys to complete and if they do complete any, they are compensated.[38] In our case, the survey company targeted people with characteristics we were interested in (managerial position, specific sectors, etc), and this was confirmed by questions in our own survey. Of course, this is intentionally not a random sample, nor representative of the general population, because we sought perspectives from specific industry actors.

The survey contained a mix of closed questions (some in Likert-scale form, others in simple yes or no form and one as an experiment) and two open-ended questions where text responses were allowed. The latter provided an opportunity for respondents to provide explanations for, or additional information related to, the responses they gave to the closed questions. Closed question responses were analysed using descriptive analysis and basic statistical analysis, with open questions analysed using abductive coding.

In total, 209 participants completed and submitted the full survey, and respondents were split relatively evenly between the two countries: 100 from Canada and 109 from the UK. There were 106 responses in the H sector (51 Canada, 55 UK), 84 in the BC sector (39 Canada, 45 UK) and 19 from the E sector (10 Canada, 9 UK). These differences in number of responses across the different industries can in part be explained by the wide variations in the number of participants in each industry. Energy has, eg by far the fewest companies in the sector.

3.2 Interview Methodology

Using the same study parameters as the survey described above (UK and Canada participants from the BC, H and energy sectors), a series of semi-structured interviews were conducted. In total, 37 interviews were conducted—20 in Canada (10 BC, 7 E and 3 H) and 17 in the UK (2 BC, 8 E and 7 H). The population

36 Shivani Vora, 'New York Hotels Make a Green Pledge' *The New York Times* (New York, 19 January 2016) <https://www.nytimes.com/2016/01/24/> accessed 7 June 2021.

37 Many survey companies provide access to survey respondents that are part of an existing large group of pre-recruited respondents (survey panel). A survey panel should not be confused with a panel survey. The latter is a method whereby the same people are surveyed at different times to understand changes in perceptions, while the former is a type of recruitment. Survey panels are particularly useful when wanting to target specific respondents which, in our case, were managerial or executive-level employees with businesses based in the UK or Canada and operating in the H, BC or E sectors).

38 Participants in these online survey panels typically are compensated either financially or through benefits such as points or travel miles, or other similar items.

interviewed were employees within these sectors who held a position that allowed them to speak to decision-making in their own company (managerial or other senior position) (9 in Canada, 4 in the UK) or representatives from trade or industry organisations who could speak about broader trends and ideas that existed overall within a sector or industry (11 in Canada, 13 in the UK).[39] Again, this study design was intentionally non-representative, but that is acceptable for this research given that the perspectives of specific actors in certain industries were sought, not the opinion of the general population, nor even the general employee base in these sectors. Other similar research uses similar methodological approaches also with great success.[40] Because of the study design, and the fact that participants self-selected, and in some cases were referred to the researchers by previous participants, we were unable to speak with an even number of participants across all sectors and both countries. While this is a limitation of the study, the data still provide concrete findings.

To gain participants, initially industry organisations in both countries were contacted with a request to send out the survey to their members[41] and this survey had as its last question, a request for participants to be part of the interview portion of the research. Unfortunately, the response rate on this survey and on the question about interviewee participation was extremely low, and therefore alternative avenues for participation were needed. Interview participation was then sought by gathering publicly available email addresses for individual companies in the three sectors and sending requests for participation. Finally, a snowballing sampling method[42] was used once the first round of interviews had been completed, with interviewees providing contact to others who might be willing to participate. Upon their completion, the interviews were transcribed verbatim by the research team and then coded using an abductive approach with NVivo.

Given the use of a semi-structured interview approach, the exact wording of questions asked to each participant varied slightly based on the responses to previous questions, but they all were drawn from the same interview themes and based on the same interview script. Initially, the participants speak about their company or organisation in an effort to put them at ease within the interview. This then moved into questions related to their thoughts on energy efficiency generally, and with regards to their perspective as a member of industry. Questions about motivations for action (general and energy efficiency related) followed. If participants had not already spoken about ideas related to norms and corporate social responsibility, then the interviewer primed them by providing examples and asking for their perspective and how they might classify these types

39 It is likely that many who self-identified as representing a trade organisation for the purposes of the interview were also currently (or had in the past been) employees of companies in the same sector.

40 Lawton (n 9) 29.

41 It is hard to know exactly how many companies received the survey because we were not given access to the mailing lists, but 16 organisations agreed that they would send out the survey link to their members.

42 A methodology where a convenience (not random) sample of the target population is relied upon to select or suggest others from the same population. See Leo A Goodman, 'Comment: On Respondent-Driven Sampling and Snowball Sampling in Hard-to-Reach Populations and Snowball Sampling Not in Hard-to-Reach Populations' (2011) 41 Sociological Methodology 347.

of actions or motivations. This specific theme was important for the broader research programme's objective. Finally, the researcher sought to get a better understanding of what methods or legal tools industry felt government could (or could not) use to help create norms. Importantly, for this article, respondents were not asked about any of Fuller's eight criteria. Their responses relating to aspects of the internal morality of law occurred unprompted, perhaps indicating even more clearly the challenges that exist.

4. ENERGY EFFICIENCY POLICY AND REGULATION IN CANADA AND THE UK

To achieve progress in energy efficiency, both Canada and the UK use a variety of instruments from the regulatory toolbox. Given that in both countries energy efficiency can be addressed through both national (UK)/federal (Canada) and devolved (UK)/provincial, territorial, (Canada) and municipal (UK and Canada) measures, and given the complexities of the regulatory structures in Canada[43] and the volume of regulation on energy efficiency,[44] we provide an overview and comparison of the two countries rather than a detailed discussion of the regulation.

In the UK, measures have been shaped by the EU's Energy Efficiency Directive.[45] In the same year as the Directive was adopted, the UK Government adopted an energy efficiency strategy[46] aimed at improving both household and corporate energy efficiency. Though updated, the strategy still informs the basis of the government's approach to energy efficiency. It identifies four barriers to greater energy efficiency and sets out measures to address these barriers which are: a small energy efficiency market, insufficient information, misaligned financial incentives (measures which do not actually target those who would benefit from the measures adopted) and undervaluing energy efficiency relative to other investment opportunities, which lead to a lack of investment in this area.

In short, a combination of command and control measures,[47] information[48] and economic incentives[49] has been used to increase energy efficiency in the UK. Some

43 You would be hard pressed to find an academic piece written about environmental law and policy in Canada which does not reference the challenges that the federalism structure, and the division of powers in the Canadian constitution bring. For a good overview, see Kathryn Harrison, *Passing the Buck: Federalism and Canadian Environmental Policy* (UBC Press 1996).

44 See the Efficiency Canada website for an example of the plethora of energy efficiency law, regulations and policy in Canada <https://www.efficiencycanada.org/> accessed 29 April 2021.

45 Council Directive 2012/27/EU of 25 October 2012 on energy efficiency and amending Directives 2009/125/EC and 2010/30/EU and repealing Directives 2004/8/EC and 2006/32/EC [2012] OJ L315/1.

46 Department of Energy and Climate Change [DECC], *Energy Efficiency Strategy: The Energy Efficiency Opportunity in the UK* (November 2012) <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/65602/6927-energy-efficiency-strategy–the-energy-efficiency.pdf> accessed 7 June 2021.

47 See, eg, the regulations regarding the sale of (energy efficient) lightbulbs for domestic use such as: Commission Regulation (EU) 2015/1428 of 25 August 2015 amending Commission Regulation (EC) No 244/2009 with regard to ecodesign requirements for non-directional household lamps and Commission Regulation (EC) No 245/2009 with regard to ecodesign requirements for fluorescent lamps without integrated ballast, for high-intensity discharge lamps, and for ballasts and luminaires able to operate such lamps and repealing Directive 2000/55/EC of the European Parliament and of the Council and

measures have been aimed at particular industries. For example, the energy industries became subject to the Energy Company Obligation, which set targets for greenhouse gas reductions, which are revised regularly through statutory instruments.[50] Similarly, energy efficiency provisions have been included in construction standards, though as Barlow[51] notes, standards have varied across time and a policy mix has been used to induce change within the sector.[52] Other industry sectors had no targets specifically focused on them, but were encouraged to adopt more efficient measures through, eg general obligations to provide information on their actions in company reports,[53] and energy efficiency labelling of appliances.[54]

The Canadian context is more diverse owing to the constitutional division of powers and the complex nature of jurisdiction over environmental law.[55] As with many areas of law in Canada, the regulation of energy is technically a shared jurisdiction between the federal government and the provinces[56] though provinces have control over many (if not most) substantive areas, especially as they relate to energy efficiency, including building codes and the regulation of public utilities as well as municipalities and their planning.[57] Given this reality, a variety of approaches have been taken across Canada to address energy efficiency, with a key similarity being the use of a range of regulatory instruments often in combination.[58] Once again, we see information combined with economic instruments and command and control provisions by the federal government.

Commission Regulation (EU) No 1194/2012 with regard to ecodesign requirements for directional lamps, light emitting diode lamps and related equipment; See also DECC, *Energy Efficient Products-Helping us Cut Energy Use* (2014) <https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/328083/Energy_efficient_products_-helping_us_to_cut_energy_use_-_publication_version_final.pdf> accessed 30th March 2021.

48 See, eg, the establishment of the Energy Saving Trust by the UK Government to Provide Impartial Advice to Consumers.

49 See, eg, the Green Deal introduced by the Energy Act 2011 and the CRC Energy Efficiency Scheme which was regulated under the CRC Energy Efficiency Scheme Order 2013, SI 2013/1119, as amended by the CRC Energy Efficiency Scheme (Amendment) Order 2014, SI 2014/502 and withdrawn by the CRC Energy Efficiency Scheme (Revocation and Savings) Order 2018, SI 2018/841.

50 For the latest iteration, see Electricity and Gas (Energy Company Obligation) Order 2018, SI 2018/1183.

51 Michael Barlow, 'Energy Efficiency in Buildings-Policy and Regulation' (2017) 29 Environmental Law and Management 211.

52 Florian Kern, Paula Kivimaa and Mari Martiskainen, 'Policy Packaging or Policy Patching? The Development of Complex Energy Efficiency Policy Mixes' (2017) 23 Energy Research & Social Science 11.

53 See, eg, the Companies Act 2006 (Strategic Report and Directors' Report) Regulations 2013, SI 2013/1970 which require the Directors' reports to address greenhouse gas emissions if their company falls within certain categories of company.

54 Council Regulation (EU) 2017/1369 of 4 July 2017 setting a framework for energy labelling and repealing Directive 2010/30/EU [2017] OJ L198/1.

55 Harrison (n 43).

56 National Research Council, *New Tools for Environmental Protection: Education, Information, and Voluntary Measures* (National Academies Press 2002).

57 Brendan Haley, James Gaede and Cassia Correa, 'The 2019 Provincial Energy Efficiency Scorecard' (Efficiency Canada c/o Carleton University 2019) <https://www.scorecard.efficiencycanada.org/wp-content/uploads/2019/11/Scorecard.pdf> accessed 31 March 2021.

58 For a full review of this, please see Haley, Gaede and Correa (n 57).

Information has been a key tool used by the federal government and is seen in the EnerGuide labelling system designed to standardise energy ratings for some consumer items across the country. It is mandatary for items such as lightbulbs, heating and cooling systems, and large appliances,[59] and voluntary for such things as vehicles and houses.[60] In relation to energy efficiency in construction, uniformity across Canada is encouraged through the National Energy Code for buildings that many provinces choose to adopt into law within their own jurisdiction.[61]

Where Canada differs from the UK is in the variety of approaches to energy efficiency across the provinces. Some provinces are leaders in energy efficiency interventions, while others seem to stand out for a specific type of intervention. The 2019 Canadian Provincial Energy Efficiency Scorecard produced by Efficiency Canada ranks British Columbia as the leader with their Energy Step Code for helping to achieve net-zero buildings alongside other interventions such as carbon pricing, capacity building for industry, transformation of the appliance and equipment market amongst others.[62] Other provinces lead on different interventions such as vehicle electrification and research and development (Quebec), building reporting and benchmarking (Ontario), focused energy efficiency programmes (Nova Scotia), minimum energy efficiency targets (Manitoba) and goals for reducing energy poverty (Price Edward Island, Nova Scotia, Manitoba).[63] What can clearly be shown by this brief overview is that both within the federal government and provinces, a mix of government interventions is being used.

While it would be possible to use these differences in regulatory approach to explain the differences in energy efficiency improvements in Canada and the UK, as the Energy Efficiency Scorecard demonstrates, the differences in regulatory approach are not so great as to preclude further discussion. In fact, both countries made use of very similar regulatory approaches, relying heavily on a mixture of information, command and control, and positive economic tools.[64] The following empirical analysis seeks to shed light on the state of shared understanding and practice around energy efficiency and in so doing to improve understanding of why regulatory interventions succeed or fail in their objectives.

59 Energy Efficiency Regulations, 2016 (SOR/2016-311).

60 National Resources Canada, 'The EnerGuide Label' (Government of Canada 2020) <https://nrcan.gc.ca/energy-efficiency/energuide-canada/energuide-label/13609> accessed 23 April 2021.

61 National Resources Canada, 'Energy Efficiency for Buildings' (Government of Canada 2020) <https://www.nrcan.gc.ca/energy-efficiency/buildings/20671> accessed 23 April 2021.

62 Haley, Gaede and Correa (n 57) 17.

63 ibid 16.

64 Economic instruments (also known as market-based instrument, financial implementation tools, etc.) can either 'encourage [actors] to undertake some activity desired by governments through the provision of financial incentives or to discourage them through the imposition of financial costs' (See Michael Howlett, *Designing Public Policies* (2nd edn, Routledge 2019) 211.) meaning they can have a positive incentive or a negative one. In the case of energy efficiency regulatory frameworks, governments have leaned towards primarily positive economic instruments, and certainly our interviewees focused on the positive incentive design when they did speak about economic instruments.

5. EVIDENCE OF GENERAL SHARED UNDERSTANDING AND PRACTICE?

The first question in applying the interactional account of law to energy efficiency regulation is 'does industry believe itself bound by an energy efficiency norm?' We know there is a plethora of law related to energy efficiency, but if industry does not see that law as having generated norms of action through shared understanding and practice, compliance may be low. We first sought to address this question through a survey which asked a very clear question about norms. In framing survey questions, we were careful to avoid academic jargon, so we referred to obligations, asking: 'Do you feel your company has an obligation to engage in energy efficiency actions although they may not be legally required?' The phrasing of 'not legally required' was used to get industry actors to consider the broadest interpretation of energy efficiency obligations. Eighty-four per cent of the survey respondents agreed that such an obligation did exist. Further, and perhaps somewhat surprisingly, these responses were consistent between the two jurisdictions (85% in Canada and 83% in the UK). These results are somewhat unexpected given the gap between these two countries in terms of their ranking on the International Energy Efficiency scorecard (UK ranking 4th and Canada 10th).[65] As such, responses to this question raise a conundrum, as they suggest that the majority of industry actors across both countries do feel they are obligated to engage in energy efficiency actions despite Canada's apparently laxer regulatory and policy standards in this area.

Turning to practice, a closer inspection of the survey responses suggests that there is little coalescence within the two countries (Table 1) and across the three sectors (Table 2). Instead, the survey data point to a variety of energy efficiency actions being taken by industry actors with some, but not a significant, variation between Canada and the UK. Similar variation can be seen if the data are further broken down to look at sector-specific actions (Table 2), with the energy sector most often being the outlier. For example, the percentage of respondents indicating energy efficiency action related to renewable energy is significantly higher in the energy sector (in both Canada and the UK) as compared to the other two sectors. Though the variation across sectors might be expected given the structural differences between them, the lack of variation between Canada and the UK does appear surprising given the different energy efficiency ratios of the two countries and the variations in regulation between the two countries discussed above.

Taken together, then, it seems that our survey results point to a potential shared *general* understanding related to an obligation for energy efficiency, but that there is little evidence in terms of shared practice around *specific* actions. It raises questions as to why the shared practice around specific actions has not arisen. Again, this disparity could be explained at a superficial level by noting the differences in regulation across and between the two countries. If, however, law depends on social practice for formulation and maintenance, as Brunnée and Toope suggest it does, it would seem that energy efficiency law is failing. Even if we see a general level of shared understanding (as demonstrated by survey responses on obligations for energy efficiency actions), more specific understandings as well as cohesion of social practice (ie what

65 Castro-Alvarez and others (n 27) 10.

Table 1Breakdown of energy efficiency actions undertaken in Canada and the UK (percentage of respondents who indicated their company did engage in action (s))

Energy Efficiency Actions	Percentage of Canada respondents ($n = 73$)	Percentage of UK respondents ($n = 85$)
EE lighting	82	74
EE appliances	60	66
Power management	56	80
Motion sensor	53	54
LCD monitors	52	53
Programmable Thermostat	49	42
Renewable energy	32	36
Leadership in Energy and Environmental Design (LEED) certified	23	13
Energy audits	19	28
Other	3	0

actors are doing to comply with this obligation or social understanding) are illusive. Perhaps, this is why, despite different energy efficiency scores on the International Scorecard and differing energy efficiency ratios, neither Canada nor the UK seems to be on track to meet energy efficiency targets and both appear to regularly introduce new approaches. We seek, therefore, to delve further into the views of industry to establish a more nuanced understanding of the barriers to developing normativity around specific regulatory actions.

6. BARRIERS TO SHARED PRACTICE

Our interviews with industry actors provide extensive insight into why the regulatory approaches seem only to be able to generate a broad level of shared understanding, rather than also being able to generate specific shared practice. The discussion of barriers to shared practice by our interviewees can be illuminated most clearly by drawing on Fuller's eight criteria for the internal morality of law as a lens. Specifically, energy efficiency regulation seems to fail very clearly to meet four of Fuller's criteria: clear and intelligible, free of contradictions, consistent and possible to obey, though these failures are not uniform across the two countries and three industry sectors of this case study. In addition, though there were fewer concerns expressed regarding the other four, what we might term 'procedural criteria', some concerns still exist with respect to these too.

We begin our discussion with the procedural criteria—law must be general, promulgated, non-retrospective and administered congruently with its intent—before turning to what we term the 'substantive criteria' of consistency, free from contradictions, clear and intelligible and possible to obey.

Table 2Breakdown of energy efficiency actions undertaken by sector and country (percentage of respondents who indicated their company did engage in some action(s))

Energy Efficiency Actions	Percentage of CA B and C ($n = 33$)	Percentage of CA H ($n = 32$)	Percentage of CA E ($n = 8$)	Percentage of UK B and C ($n = 34$)	Percentage of UK H ($n = 43$)	Percentage of UK E ($n = 8$)
EE lighting	82	88	63	82	70	63
EE appliances	52	63	88	62	67	75
Power management	55	59	50	82	81	63
Motion sensor	52	53	63	53	53	63
LCD monitors	45	56	63	56	47	75
Programmable Thermostat	48	47	63	44	40	50
Renewable energy	30	25	63	53	16	75
LEED certified	21	22	38	18	9	13
Energy audits	18	19	25	21	33	38
Other	3	3				

6.1. Procedural Criteria

6.1.1 General

The need for laws to meet Fuller's criterion of being generally applicable may at first appear to be a challenge in the context of energy efficiency. While some regulations, such as those regulating the sale of types of lighting, may be applicable to all members of the public and all industrial sectors, many of the regulations do focus on particular sectors. They are, however, generally applicable within those sectors and this ought to be sufficient to meet Fuller's criterion. Indeed, only four of our interviewees raised the generality of regulations as an issue. Two discussed, in a relatively abstract sense, the need for problems to be tackled at a general level. The other two interviewees who spoke to this issue commented on a perception that energy efficiency regulations (or indeed regulations in general) may impact some parts of the sector more than others. A Canadian interviewee spoke about the difference between the home building side of construction and other areas noting that '. . .there's obviously you know our side of the industry is heavily regulated versus the home building side which is you know anybody who grabs a $4 hammer at Canadian Tire can be a framer in house building' (CANBC1). Similar distinctions between different types of BC were raised as a concern by one UK respondent who noted '[i]f you're building wooden homes, a lot of those timber frame homes, those people feel the . . . calculations don't really offer the same sort of protection in the planning process, so we've got actually a big tougher for them for meet the . . . regulations.' (UKBC2). Given the relatively small number of comments that related to Fuller's criterion of

generality, however, on the whole, it does not appear that there are significant concerns with a failure to meet this criterion.

6.1.2 Promulgated

It might generally be assumed that the question of whether law has been promulgated will always be answered in the affirmative. The practice of publishing laws through official publications is long settled and the use of government web pages increases visibility of legislation and regulations to the public (at least in theory.) Despite this, four of our respondents made it abundantly clear that making energy efficiency law and policy available to the public through these official channels may not be enough to really satisfy the promulgation criterion, with a fifth also making comments on the importance of promulgating the law. One respondent (in reference to incentives for solar thermal water heater installation) said 'Nobody knew about it. So they were giving out grants, nobody knew that that was a thing. So maybe one person got it installed' (CANE3). A second respondent (UKE6) raised a slightly different point, speaking of the need to give advance information on government policies. They spoke of the difficulty of planning regarding training and recruitment of staff due to the fact that government policies were not publicised sufficiently far in advance. Again, the small number of interviewees raising these points suggests that there is no pattern of failure in ensuring that laws are promulgated, though there may be ways to improve on any perceived failures. Of course, an alternative explanation for why this criterion does not seem to be a significant factor may be that industry actors do not realise that they are unaware of energy efficiency laws.

6.1.3 Non-retrospective

That laws should not be retrospective is not something we find respondents in our study mentioned much (only three interviewees made comments germane to this issue.) It therefore does not seem to be a major cause of the failure to develop norms for shared understanding and practice around industry energy efficiency actions. In each case where an interviewee did make a comment, they appeared to indicate that some measures are, in effect, rewarding actions that have already been taken, rather than regulating future action. As one noted, 'it shouldn't be about paying people who have already done all the hard work' (CANBC5). It appears that the interviewee demonstrates a good grasp of the non-retrospective criterion, but as only three interviewees raised this issue, we would suggest this is not a major challenge with regards to energy efficiency laws.

6.1.4 Administered congruently with its intent

The requirement that government officials administer the law congruently with its intent was only raised by four of our interviewees. One noted the need for more enforcement in the context of health and safety regulations (something used by our interviewees as an example of an idea which developed into norms of practice within the sectors). This respondent stated '. . . you know a lot want more regulations a lot want more uhh oversight and a lot more uhh penalty if you are not compliant in our side of the business so over and above that I mean . . . they want their people to go home safe' (CANBC10). Another spoke more positively to the fact that governments

do seek to ensure that laws are implemented as they were intended to be, noting: 'I am not aware of governments trying to cut corners' (UKE3). We could guess at the reasons interviewees did not raise congruence as an issue. Perhaps, it is because many of the regulatory instruments used in relation to energy efficiency are economic instruments and so enforcement, eg is not a significant issue. Perhaps it is because where enforcement action is a possibility, government officials do act in accordance with the law, or where they fail to do so, it is to the advantage of industry. While we cannot provide concrete conclusions on this point, we suggest that the relative lack of comment by our interviewees indicates that, on the whole, energy efficiency regulations are administered congruently with their intent.

6.2 Substantive Criteria

6.2.1 Clear and intelligible

The very notion that law must be clear and intelligible seems obvious, and yet 11 interviewees (30%) indicated that the law, or how to comply with it, was either not clear or not intelligible. Interestingly, eight of these interviewees were from the Canadian Building and Construction Sector (which accounts for 80% of that category of interviewee) and three were from the UK Energy Sector (almost 40% of that category of interviewee). Although one UK interviewee noted that they thought 'people have plenty of information and I don't think the reason that they are not coming on stronger is because they are not aware of it [rooftop solar panels] or they don't understand the benefits. I think the people who know about it would do it if it were cheaper, or if they had to' (UKE1), that interviewee appears to be somewhat of an outlier.

A representative comment came from CANBC1 who, when speaking about a tax credit programme, indicated that 'I think a lot of people really didn't understand what it [the tax credit] was and that was helping to pay for the development of uhh innovating methodologies.' Seven other respondents spoke specifically to the idea that alongside new (or existing) energy efficiency law, there must also be education and training provided as a way to ensure better compliance. One respondent said '. . . any province that's advocating any new or higher energy efficiency that does not have mandatory education and training is irresponsible' (CANBC2), while another noted that '. . . once senior leaders, CEO of companies understand the business value of it [CSR], they take ownership, and they understand, they support it.' (CANBC6) Further, a respondent noted that 'I think there needs to be really strong central government commitment that says, this is what we expect of our buildings. You know, setting the scene for 2030 and beyond, that says, the buildings we've got at the moment are not acceptable, they've got to be different.' (UKE7) Taken together, this demonstrates that many in these two sectors feel that more information and education to all (from trades to building managers to the C-suite[66]) would improve compliance. One could argue that the provision of information and education is a separate issue to clarity in the law—the law could be technically clear, but industry may still require education as to its content and implications. However, as Fuller

66 See Boris Groysberg, L Kevin Kelly and Bryan MacDonald, 'The New Path to the C-Suite' (2011) 89 Harvard Business Review 60.

notes, '[t]he desideratum of clarity represents one of the most essential ingredients of legality,'[67] and the implication here is that, right now, this does not exist in the minds of the regulated. In other words, focussing on legal clarity alone when considering whether laws are clear or not may be one of the causes of regulatory failure. What may be required is greater clarity regarding expectations or opportunities for the regulated.

The second issue that arose in our interviewee responses is that the range of regulation, eg some focussing on energy efficient lighting, others on solar panels, others on reporting requirements, left interviewees unclear on where they should focus their particular attention. In other words, although individual regulations such as the focus on energy efficient lighting might reflect specific aims, the variety of foci leaves those targeted by energy efficiency regulation somewhat unclear on what actions they should actually be taking. As one respondent put it 'even just knowing . . . what is expected of me. . .And giving building owners guidance on those long-term expectations so they can start planning' (CANBC5). As Fuller himself put it '[i]t is obvious that obscure and incoherent legislation can make legality unattainable by anyone, or at least unattainable without an unauthorised revision which itself impairs legality.'[68] Additional law is not necessarily the solution to this, rather it points again to the need for clear information or education to ensure the regulated understand what the law requires and how to prioritise actions.

6.2.2 Free of contradictions

There is a need for energy efficiency laws to be free of contradictions with one another and with other law that is necessary in order to progress towards shared action. As Fuller points out, 'legislative carelessness about the jibe of statutes with one another can be very hurtful to legality.'[69] In total, 11 respondents raised this issue in their interviews. How they raised it varied, eg two spoke very clearly about the need to link the regulation, or policy responses to energy efficiency to policy in other areas such as, those addressing pollution from run off, or protecting biodiversity (UKBC2). A couple spoke to the fact that contradictions between different areas of law prevented the take up of some energy efficient initiatives. For example, UKE4 spoke to the fact that air quality requirements stopped the uptake of biomass boilers which had otherwise been promoted as a more sustainable fuel source. Two further interviewees focused on complementarity rather than contradictions. For example, UKBC1 spoke to linking governmental interventions together as being particularly important, stating that '. . .this is the policy and if you want some funding then you follow the policy. . .. Whether it's regulated or whether it's just a policy drive if you want the money then you need to do this.' Although the way the interviewees raised the issue of regulation being free from contradictions varied, comments by respondents that either this criterion is important to comply with if regulation is to be effective, or that there are problems in practice tells us that, in the views of the regulated at least, there is some work to do to ensure that regulations are free of contradictions.

67 Fuller, *The Morality of Law* (n 22) 63.
68 ibid.
69 ibid 9.

Some of the comments, as indicated above, also tell us very clearly that where contradictions arise, the likelihood of regulations successfully achieving their intended goals may be reduced. This is not an entirely new finding, given the literature on policy mixes and regulatory toolboxes,[70] but the fact that it is still being raised as an issue tells us that more is needed to ensure, at the very least, that the regulated understand how regulatory systems fit together.

6.2.3 Reasonably consistent (across time and geography)

Our interviewees frequently raised the problem of (in)consistency in the context of energy efficiency regulation (15 (41%) interviewees spoke to this issue). This came across most clearly when respondents spoke about incentive programmes, often related to new technologies.[71] A general feeling of frustration and disappointment was expressed as to how economic instruments (often those providing incentives such as grants or subsidies) had been implemented. For example, both respondents from the UK BC sector spoke of challenges caused by inconsistency in government approaches to energy efficiency. As one said: 'how could they, ah, meet that standard if they've got to wait 6 years? That standard could change, but we saw that with energy efficiency and we saw that with renewables, um, a lot of the subsidies are taken away by the government right in the middle of the consultation period on renewables' (UKBC2). They continued by saying that it is the small- and medium-sized enterprises in their industry that are particularly impacted by 'moving the goal post. . . and that's a major problem in terms of energy efficiency, or even let's say just improving the building stock' (UK BC2).

Interviewees from other sectors also raised the problems caused by a lack of consistency. One interviewee, speaking of subsidies for solar panels and the role they had in terms of generating installation and manufacturing and general adoption suggested 'they [the government] definitely pulled out way too soon and they did not establish the economies of scale that would make it viable subsidy free' (UKE1). Another respondent spoke about tax breaks saying '. . . there are tax breaks to do it that the government have offered and as soon as they offer them and people take them up they reduce them, so it doesn't encourage anybody else to do it' (UKH7). In addition, one respondent provided an example of a period in time (2007/08–2011), where local performance indicators were linked to funding, indicating that

70 The idea of the best 'fit' and optimal policy mixes in terms of complementarity and contradictory legal instruments is something pioneered in many senses by Gunningham and Sinclair's 1999 work (See Neil Gunningham and Darren Sinclair, 'Regulatory Pluralism: Designing Policy Mixes for Environmental Protection', *Environmental Law* (Routledge 2019). While there seemed to be period of time where other areas of legal instrument and policy tool research were the emphasis (see, for an overview, Michael Howlett, 'From the 'Old'to the 'New'Policy Design: Design Thinking beyond Markets and Collaborative Governance' (2014) 47 Policy sciences 187), the focus has again turned back to understanding instrument interactions (eg Pablo Del Río, 'Analysing the Interactions between Renewable Energy Promotion and Energy Efficiency Support Schemes: The Impact of Different Instruments and Design Elements' (2010) 38 Energy Policy 4978; Yilin Hou and Gene A Brewer, 'Substitution and Supplementation between Cofunctional Policy Instruments: Evidence from State Budget Stabilization Practices' (2010) 70 Public Administration Review 914; Pablo Del Rio, 'On Evaluating Success in Complex Policy Mixes: The Case of Renewable Energy Support Schemes' (2014) 47 Policy Sciences 267.)

71 A fuller discussion of this and the subsequent seemingly negative perspective that this led them to have on positive economic instruments (especially by those in the UK) is contained in Besco and Kirk (n 8).

this idea was in place for a 3- to 4-year timeframe and then disbanded or replaced (UKE7). While some Canadian respondents had similar sentiments, their comments tended to be more broadly focused on misalignment of incentives and payback periods rather than on (in)consistency of the regulation across time. That the majority of responses highlighting inconsistency across time were from the UK is perhaps not surprising, given what others have published related to some of these interventions. For example, Bright and Weatherall note the failures of the Green Deal and its relatively short life.[72]

There was also evidence presented by respondents about challenges in terms of consistency across geography, an important issue with regards to the adoption of new products. As one Canadian respondent put it (when talking about energy efficiency requirements under a provincial building code) '[w]indows and doors, most of the manufacturers products don't meet uhm the SB10 requirements. . . for. energy efficiency and they say well we're not going you know completely change and create a customize product for . . . until everybody gets on board with it' (CANBC7). The implementation of prompt payment legislation in BC in Canada was also used as an example by interviewee three from the Canadian BC sector, in terms of inconsistent implementation across provinces. In saying '[s]ome of the things they could do we're already moving towards, prompt payment legislation which is happening in Ontario, is moving forward in Winnipeg and Nova Scotia. In Saskatchewan we have umm a written commitment that they're going to be introducing in the fall here' (CANBC3), this respondent demonstrates the differences in implementation, though over time it seems this type of law is becoming more common.

Inconsistent standards and requirements across different provinces and territories inevitably lead to challenges with compliance and implementation. Ideas of consistency across geography also came out of interviews in the UK, with particular focus on the relationship between EU requirements and the UK implementation (here in reference to Energy Performance Certificates (EPC)), '[b]ecause we had to upgrade the EPC, so the EPC are based on European directives that the UK then put them on steroids, I mean they're way better than the others. A long way ahead of other countries' (UKE8). There was also some reference to inconsistencies arising because of the challenges raised by split levels of responsibility for energy efficiency. As one UK interviewee indicated '. . .it's a little bit fragmented especially in the UK because it's the central government which is on some levels, but actually, the local authorities, if you've got the planning process to local authorities, and the plan making process which is actually a little bit different than the planning process' (UKBC2). In other words, they were pointing to the fact that the variations in planning processes may lead to different energy efficiency standards in effect being in place across the UK. A further UK respondent spoke about variations in energy efficiency requirements for new-build, indicating that the requirements across the country varied when they stated 'you are seeing that on some local authorities, but if that became like a national policy that would be amazing' (UKE1).

72 Susan Bright and David Weatherall, 'Framing and Mapping the Governance Barriers to Energy Upgrades in Flats' (2017) 29 Journal of Environmental Law 203.

Consistency in regulation across time, geography and regulatory competencies is clearly an issue that requires further attention. We would suggest that these issues are ripe for further research to develop understanding of the bounds of time, geography and competencies that need be considered in developing regulation that is sufficiently consistent to meet Fuller's criterion.

6.2.4 Possible to obey

Fifteen interviewees made comments which spoke of the importance of Fuller's 'possible to obey' criterion. The main component highlighted by interviewees in this context, related to capacity, though some also spoke to financial feasibility. To the first point, one respondent in the Canadian BC sector said, eg 'a lot of it is then equipping, because we have this massive green gap right now in terms of the design standards and then how we can actually build these buildings' (CANBC4). In other words, getting from standards to actual implementation in buildings is a huge hurdle which requires specialised training because of new procedures or technologies which means, as a second respondent in the same sector suggested '. . . most of the people in construction are not skilled to do the work that's required' (CANBC5). In other words, even if the law is understood, the workforce capacity to carry it out is not currently available.

The financial costs of responding to calls for increased energy efficiency action also presented a challenge to obeying regulations for some in the industry. They raised the problem created by the payback period associated with taking the necessary action and noted that this arises even in the context of the application of grant or incentive programmes. One respondent presented the issue through the example of solar panels '. . . until prices start to come down as these items become more popular. . . and you know as an example you give solar panels or photovoltaic panels or these sorts of things you know it's really hard in our location and climate to uhm get a reasonable payback for clients to implement those on their projects' (CANBC7). Another respondent raised these issues in the context of the need for deep building retrofits to tackle climate change noting that '. . .most . . . will not payback from energy saving, umm directly under the current framework with the current technology pricing' (CANBC5). They went on to note the need to and challenges for regulators with offering 'incentive programmes, which essentially incentivize people to do the low hanging fruit, the stuff that does have a payback, or close enough to a payback that the marginal incentive would tip it over'. Here they noted that the regulators would be 'paying for a lot of free ridership essentially' (CANBC5). In effect, this interviewee is speaking of disconnect between targets and overall goals and the support, or incentives, provided by the regulatory framework leading some actors to feel unable to comply with (or obey) the regulation. The same respondent spoke of the difference in ability to engage with programmes and policies that exists between different sizes of corporations and that this might be one of the challenges with engagement.[73]

73 See Stephen Brammer, Stefan Hoejmose and Kerry Marchant, 'Environmental Management in SME s in the UK: Practices, Pressures and Perceived Benefits' (2012) 21 Business Strategy and the Environment 423.The article provides an excellent look at small versus medium sized enterprises reasons and barriers

A further issue raised by five interviewees is the difficulty of processing the volume of regulation on the topic. It seems that this too can act as a barrier to their ability to comply with (or obey) the regulations. For example, one respondent spoke about 'put[ting] together a document whereby we reviewed all the current regulations and guidelines and we tried to put them into one document to understand and for everyone to interpret' (UKBC2). While this does not necessarily indicate there are too many rules, it does suggest that there are enough that actors might struggle to know them all and therefore they find complying with, or obeying those laws problematic. Another respondent commented on the problems posed by regulatory processes noting that they had 'been working with the ministry of environment on coming up with best management practices as opposed to having to get a permit all the time' (CANE7). The comment indicates that the need to continually apply for permits is problematic. A similar issue was raised by a respondent who spoke about a new approach being taken in British Columbia called the Step Code. They suggested that focusing on this on its own is a good idea noting 'from what I'm gathering, umm there's been a lot of umm attention placed on that [Step Code] and just you know like let's strip away all this other stuff, even some of the other programmes and initiatives we would run and just like, ok how do we get down to umm getting basically prepared for the BC Step Code.' (CANBC4). In the same way, a further respondent spoke about identifying equivalents to minimise the number of reporting frameworks that must be used, in so doing they were, perhaps, suggesting ways to reduce regulatory burden: 'But if you do have your own requirements, it would be good to have alternative compliance paths, or look at programmes that have equivalents. So that you don't have to go through two different reporting frameworks' (CANBC5). While this does not necessarily point to the law being impossible to obey, it does point to practical challenges that stand in the way of effective compliance with the law. In the context of improving energy efficiency, effective compliance (ie compliance which really does lead to improvements in the rates of energy efficiency) really does matter if we are to meet our goals for increased energy efficiency. That the regulatory burden might interfere with the possibility of obeying or complying with regulations should not be a surprising finding as the idea of regulatory burden on industry is not a new problem.[74] As one respondent noted 'regulatory burden is something that everyone tries to avoid as much as possible' (CANE1).

What we are seeing then is the regulated noting that even if, in theory, the regulations are possible to obey, there are many practical problems that make them

to engagement with environmental actions. Furthermore, it is fairly well known that one common challenge for Small and Medium Sized Enterprises (SMEs) is their limited resources—financial, human and material—especially as compared to larger corporations (See Rodney McAdam, 'Large Scale Innovation Reengineering Methodology in SMEs: Positivistic and Phenomenological Approaches' (2002) 20 International Small Business Journal 33.).

74 See, eg, Thomas D Hopkins, 'An Assessment of Cross-National Regulatory Burden Comparisons' (2005) 33 Fordham Urban Law Journal 1139; Organization for Economic Co-operation and Development, *Overcoming Barriers to Administrative Simplification Strategies: Guidance for Policy Makers* (OECD Publishing 2009). See also Antoine Dechezleprêtre and Misato Sato, 'The Impacts of Environmental Regulations on Competitiveness' (2017) 11 Review of Environmental Economics and Policy 183. We have even started to see specialized consultancy firms pop up to help deal with regulatory compliance for environmental law.

challenging to comply with. While it would be possible to argue that all the issues raised by the interviewees are problems for the regulated to resolve, they also point to the need to consider the capacity of industry to act. If industry actors simply cannot respond to the regulation as it is intended, it may be that no practice of legality emerges. Some of the issues that the interviewees raised could again be addressed through the provision of information, education and training programmes; however, it is clear that more may be required in terms of financial support to industry actors, and consideration as to what support is provided to corporations of different sizes.

7. CONCLUSIONS AND RECOMMENDATIONS

Our paper has demonstrated that refocussing attention on normativity in the context of regulation of industry actions the context of diffuse issues, such as energy efficiency, generates new understandings of the causes of regulatory failure and potential locations for improvement. As such, it demonstrates the utility of applying the interactional account of law to domestic regulation.

Our empirical study reveals that industry actors do not always act in a purely rational way, and, moreover, are not always able to act as rational actors even where they might wish to do so. That is, even if they are aware that there are laws and obligations regulating energy efficiency actions (which both our survey responses indicate and the lack of concern over promulgation demonstrates), there is inconsistency of follow-through and action. That lack of follow-through seems odd, if, as rational actors, they ought to be concerned to avoid penalties and to take advantage of incentives. In scrutinising whether or not a shared understanding exists, we found that while industry actors did have a shared understanding regarding a general energy efficiency norm, shared practice proved to be relatively illusive. Instead, we found that a range of measures were taken by industry actors. It is likely significant that the most obvious coalescence of practice and understanding is found in respect of the use of energy efficiency lighting, the one energy efficiency action largely controlled through clear, simple and consistent command and control regulation. Other (potential) actions saw far less coalescence in the practice of the regulated. These included the use of energy efficiency appliances, or adoption of power management, the types of actions that are encouraged through information instruments such as labelling (appliances) or corporate reporting, rather than being mandated. There was also less coalescence around renewable energy generation and audits, actions often encouraged through positive economic instruments. Our analysis in this article did not, however, address the suitability of different instruments to generate normativity (we have addressed this issue in a separate article).[75] What was drawn out in this article is that, in the eyes of the regulated, laws seemed often to fall short of Fuller's criteria for the internal morality of law and we suggest that it is this that stops the industry actors from fully implementing measures to comply with the energy efficiency obligation and so stops shared practices fully emerging. In other words, it appears critical to norm formation that Fuller's eight criteria are all complied with, without that energy efficiency regulation may fail to motivate the anticipated industry action.

75 See Besco and Kirk (n 8).

At a general level, the findings appear largely consistent across the two countries and three sectors in our empirical study, with some notable exceptions. It seems that issues with the clarity and intelligibility of energy efficiency law plague the Canadian BC sector and to some extent the UK energy sector, while this was not something raised by others. Further, while consistency of law is an issue that was raised across both countries and all three sectors, the focus on consistency in terms of time (and legal tools changing frequently) appeared a much more significant challenge to industry in the UK than in Canada. Canadian respondents spoke more about consistency across geography, which may not be surprising given the greater jurisdictional complexities inherent in Canadian environmental law. Finally, hospitality respondents did not make many statements at all that were clearly connected to Fuller's criteria and energy efficiency law. Perhaps, this is because, of the three sectors, hospitality is probably the one that to date has been targeted least directly by energy efficiency frameworks, though there are a multitude of other possible reasons. This is surely something for further investigation. Even with these points of divergence, neither Canada nor the UK's energy efficiency regulatory framework is entirely successful at meeting Fuller's criteria. This may be why, despite the significant difference in their current rates of energy efficiency and the differences in regulation in the countries, neither is seeing shared practices for energy efficiency and norms of action develop to the degree required to meet their energy efficiency goals.

Our findings demonstrate failures in respect of each of Fuller's criteria; however, for some criteria, those we term procedural criteria—generally applicable, promulgated, non-retrospective and administered congruently with intent—only a small number (3–5) of interviewees raised comments. We cannot, based on these findings alone, point to these as significant issues to be addressed through regulatory reform. That leaves four criteria, the substantive criteria, where there appeared to be failures across the regulatory systems: free from contradictions and clear and intelligible (both commented on by 11 interviewees) and consistency, and possible to obey (both commented on by 15 interviewees). While it may be possible to discount some of these issues given that the percentages of interviewees which commented on them are not particularly high, given that our study participants are likely more engaged or interested in energy efficiency than others within their industry sectors,[76] it seems likely that if the interviews were extended to industry actors which were not actively engaged or interested in energy efficiency there would be less acknowledgement of the existence of a general energy efficiency obligation, and more problems evident in relation to each of Fuller's criteria. Moreover, when viewed in the context of the norm lifecycle,[77] the number of interviewees raising these issues becomes

76 This gets at the inherent bias in the sampling we used and the fact that our findings somewhat underplay the factors that impact on the generation of a norm of energy efficiency. Our survey participants and interviewees, volunteered to participate in a project on energy efficiency, or had been recommended as interviewees on such a project, and therefore are likely more interested/engaged in energy efficiency than are others within their industry sector. Given that we still find significant gaps in terms of implementation of regulatory obligations and evidence of shared understanding and practice amongst our (relatively engaged) participants, our finding are probably quite conservative.

77 See, eg, Martha Finnemore and Kathryn Sikkink, 'International Norm Dynamics and Political Change' [1998] International Organization 887; Bastin Tony Roy Savarimuthu and Stephen Cranefield, 'A Categorization of Simulation Works on Norms', *Dagstuhl Seminar Proceedings* (Schloss Dagstuhl-Leibniz-

more significant. The norm lifecycle sees norms emerge, reach a tipping point (generally accepted as arising at the point at which 25–40% of actors adopting an action)[78] and then become more generally adopted through a norm cascade. As Sunstein and others[79] demonstrate, however, small changes may cause a norm to fail either to emerge or to remain constant. In this context, it must be of significance that almost 30% of those interviewed raised issues with the fit between energy efficiency and the clear and intelligible and free from contradictions criteria as issues and over 40% with the consistency and possibility to obey criteria. These figures signal that there is a risk not only that industry responses to the regulation will not match the regulatory goals, but that their responses could be copied by others leading to a new norm of behaviour being followed by industry actors. If that new norm of behaviour does not fit with the regulatory goals, regulatory failure will occur.

With regards to the specific findings, it is notable that our findings indicate that, regardless of whether or not regulation was technically clear and intelligible to those within the legal profession, it was often viewed as not being clear and intelligible to the regulated. In part, this was due to the need for information and education to be attached to the introduction of new regulation, something that is at least recognised as an important part of the regulatory process. Perhaps, the more significant finding is that the regulated needed information on how to prioritise the range of demands placed upon them by regulation aimed at improving energy efficiency. This finding appears to be linked to the challenges 30% of interviewees raised with regards to law being free of contradiction. In this context, interviewees spoke to the need for complementarity in law and some highlighted conflicting policy goals and regulations, which left them having to prioritise which to respond to. These two issues are also likely linked to the challenges 41% of interviewees highlighted relating to the (im)possibility of obeying the law. There they highlighted financial challenges, alongside the challenges of acquiring a suitably skilled workforce to implement energy efficiency measures. Given that each company (particularly small- and medium-sized enterprises) will have limited resources with which to address energy efficiency measures, the importance of prioritisation becomes key. It may also explain the range we see in practical measures taken—corporations have, perhaps, chosen to prioritise the measures that are of greatest benefit, or least cost to them, or to prioritise those that are most attractive to their customers. Industry priorities may not, however, coincide with priority actions to improve energy efficiency at the national level, nor with predictions of how rational actors would respond to the regulatory imperatives. For example, grants to install energy efficient equipment, which appear appealing to industry actors when viewed through the rational actor lens, may, in practice, be foregone if the corporation lacks the skilled workforce to install, or use the new equipment, or if they are sceptical of whether the financial incentive will remain long

Zentrum für Informatik 2009); Emma Sjöström, 'Shareholders as Norm Entrepreneurs for Corporate Social Responsibility' (2010) 94 Journal of Business Ethics 177; Mahmoud and others (n 5).

78 See, eg, Malcolm Gladwell, *The Tipping Point: How Little Things Can Make a Big Difference* (Little, Brown 2006).

79 See, eg, Cass Sunstein, *How Change Happens* (MIT Press 2019); Raymond H Brescia, 'On Tipping Points and Nudges: Review of Cass Sunstein's How Change Happens' (2020) 34 Notre Dame Journal of Law, Ethics & Public Policy 55; Lawton (n 9).

enough for them to take full advantage. This latter point connects to the third criterion where energy efficiency regulation regularly appeared to fall short: in its consistency of application across time or geographical location. Frequent changes, in particular, in relation to economic incentives to adopt new technology, reduced the willingness of industry to engage with these instruments. So too, differences in standards across geographical locations in which corporations acted reduced their willingness to invest in new practices or technologies.

Each of these problems raises similar issues. In each, the primary problem appears to be that companies lack the resources to respond to (all) the regulations being implemented. Those resources may be in the form of information, skilled workforce or financial capacity to act. The first two of these issues could be addressed through the provision of more, or better targeted information and education at no, or minimal cost to industry actors. One possible solution proposed by five respondents related to the use of centres of innovation (or something similar). Centres of innovation were presented by the respondents as locations where interested individuals could learn about new ideas, what leaders in the field are doing and 'go talk to those people [engineers, architects, contractors, trades] and find out, you know, what are the most innovative things going on in construction' (CANBC3). It is, however, also possible that such centres could be used as a vector to provide targeted communication to relevant sectors of industry with regards to the introduction of new regulations alongside information or education programmes designed to facilitate positive implementation of, or compliance with the regulation.

Financial capacity to act may be harder to rectify on its own; however, we suggest that the provision of information and education may go some way in rectifying this challenge as may addressing concerns regarding consistency of regulation across time and geographical location. This leads to a further issue that arises across the substantive criteria: the significance of time. Time may of itself be a resource that is in short supply, in which case the industry actor's solution may be to invest in enhancement of their workforce, or equipment. This of course leads back to the question of resources, discussed above. Here, however, we focus on a different aspect of time. In short, if responding to new regulation requires industry actors to make a financial commitment, whether through purchasing a new appliance, or investing in training for staff or hiring additional staff, then they need sufficient time in which to make that investment and to secure a sufficient financial return for it. One would assume that this point should be obvious to those viewing regulation of industry actors through the rational actor lens. The frequency with which the issue was raised by our interviewees suggests, however, that, at the very least, practice is not coinciding with the theory of what motivates industry actors to act. Our research was not designed to and did not reveal the optimum length of time for regulation to be effective, and indeed the issue is too complex to provide a single solution applicable to all regulation. We suggest, however, that further research into the timeframes for regulatory effectiveness is urgently needed and that such research should consider the impact of information and education interventions as mechanisms to intensify the speed at which full implementation or compliance is achieved. Indeed, given the focus on norms within this article, we suggest it may be appropriate to build on the findings in this article and insights

from the literature on norm lifecycles, to identify more precisely where sticking points in the generation of norms occur.

We further suggest that these shared issues—the resource demands of responding to regulation and the significance of time—are likely germane to many of our current sustainability challenges. Climate change, eg throws into stark relief the conflicting demands of time (the need for urgent action) and question of corporate capacity for action across corporations of all sizes and in all industries. We suggest too that our proposals regarding the need for improved vectors to disseminate information and education, along with the need for further understanding the timeframes required for regulatory effectiveness and potential sticking points in the norm lifecycle apply equally in these areas.

ACKNOWLEDGEMENTS

This project was funded by a Social Sciences and Humanities Research Council of Canada Grant (430-2016-00392). The authors would like to thank Kelsey Koebel, Lindsay Lucato, Christina Gagliano-Veiga and Jae Page for their valuable contributions as research assistants on this project. All errors or omissions rest with the authors.

Conflict of interest statement. None declared.

INSTITUTE FOR INFORMATION LAW

Creative Commons Licenses Legal Pitfalls: Incompatibilities and Solutions

Melanie Dulong de Rosnay

Institute for Information Law
University of Amsterdam

Version 1.0
1 September 2010

TABLE OF CONTENTS

Executive summary

Creative Commons licenses have been designed to facilitate the use and reuse of creative works by granting some permissions in advance. However, the system is complex with a multiplicity of licenses options, formats and versions available, including translations into different languages and adaptation to specific legislations towards versions which are declared compatible among each other after an international porting process. It should be assessed whether all ported licenses cover exactly the same subject matter, rights and restrictions or whether small language differences may have an impact on the rights actually granted and legal security of current users or the availability of works for future generations to access and build upon.
Besides, other possible sources of legal uncertainty and incompatibility, as well as their actual or potential consequences, need to be evaluated, such as the validity and enforceability of the licenses across jurisdictions with different and possibly inconsistent legislations, the variations between the licenses summary and the licenses text written in legal language, the interoperability with other copyleft licenses.
This study presents the different licenses (chapter 2), identifies various possible sources of legal incompatibility (chapter 3), evaluates their actual impact (chapter 4) and finally proposes options to mitigate risks and improve compatibility, consistency, clarity and legal security (chapter 5).

Many options are proposed to answer to users' needs. But besides creating incompatibilities among works licensed under different options, this choice has information and political costs. Reducing the number of options would lead to a clearer definition of freedom, make the choice easier for users and diminish incompatibilities between works licensed under different options.
An analysis of the licenses clauses allows finding out what is exactly covered and whether it is made clear to the user, which is necessary to provide legal certainty and security. The license elements, which are very visible, may be hiding the substance of the license to the user, who has to read the main clauses behind the options. Besides information costs, the question is whether these main clauses are not only visible, but also clear substantially to the user. Knowing precisely which rights are granted by whom on which subject matter is essential for the validity and the coherence of the system.
A systematic description of the main provisions of the eight clauses of a CC license in its unported 3.0 version will allow clarifying what is the subject matter. Then, comparing the core grant of the 3.0 unported license legal deed with the other licenses versions, jurisdictions and formats, will allow identifying actual differences and potential sources of incompatibilities. Most of the core grant is not mentioned in the Commons Deed and therefore not very accessible to the average user, who is nevertheless expected to consent to the legal code.

Before analyzing the compatibility between licenses, the compatibility with international law will be checked in order to detect possible inconsistencies or confirm that the system is viable. Obviously, the licenses do not have to mention all the notions of the international conventions and can go beyond, but it will be checked which notions are exactly covered in order to make

sure that no right or party has been left out. Indeed, the grant intends to be as broad as possible and it can therefore be expected that all works and all rights are addressed by the licenses and that no restrictions on the nature of works and rights covered are hidden behind the long wording. Scrutinizing the licenses' optional elements and main clauses will allow detecting a few formal inconsistencies to be fixed.
After examining how the licenses clauses are compatible with copyright law, it will be considered whether the licenses as a whole are compatible with contract law and consumer law. Analyzing the legal nature of the Creative Commons tools, being contracts or licenses, allows identifying possible incompatibilities with applicable law. It should be verified whether the agreement is valid and if consent between parties can be reached. If the agreement is deemed invalid and consent has not been reached after all, permission will not deemed to have been granted. Licensors may not be able to request the enforcement of non-copyright infringement related conditions even if they apply to acts triggered by the exercise of a copyright-related right, and licensees might not be able to claim the exercise of rights beyond copyright law which is fully applicable by default, and thus reproduce the work freely. Finally, specific attention will be dedicated to the Share Alike clause reciprocal effects and the transmission of obligations to third parties which should be bound by the conditions. Indeed, the system would not be sustainable if the agreement enforceability would stop after the first round.

After studying the licenses clauses and their possible incompatibilities with copyright and contract law, the issue of licenses proliferation and internal incompatibilities within the system will be studied. Two sources of differences are visible from the license interface (formats and options) but actually five sources of differences between the licenses may raise incompatibilities issues:
1. The licenses formats, the machine-readable code, the human-readable common deed and the legal code (formats), it could be possible that a licensee is not aware of important limitations which are available only in the middle of the legal code.
2. The licenses different options and combinations: BY, BY-SA, BY-NC, BY-ND, BY-NC-SA, BY-NC-ND (options), making it impossible to remix works licensed under incompatible options.
3. The licenses successive versions: 1.0, 2.0, 2.5, 3.0 (incremental versions), it will be analyzed whether the differences between the successive versions create incompatibilities between licenses carrying the same optional elements.
4. The differences between the licenses adaptations to various jurisdictions, the porting process has been engaged for the six combinations and at least one version for over 50 countries or jurisdictions (jurisdiction or ported versions). The Share Alike clause admits the relicensing of an Adaptation under a license from another jurisdiction. They are declared compatible, but are they really compatible, do they cover the same subject-matter, offer the same scope of rights and contain the same limitations? The goal of the international porting process is to facilitate local implementation, avoid interpretation problems and improve compatibility with copyright legislations. But it may actually lead to a contract law problem, because a Licensor is expected to consent to the Adaptation of her Work to be licensed under different, future, unidentified terms. This study does not analyze and compare all the 50 versions, but uses some selected examples to evaluate the contamination risk which may occur from the first generation of derivative works, and grow exponentially after several generations. Examples include provisions related to the limited warranties and representation, moral rights, the inclusion of related and databases rights in the definition of Work, the scope of applicable rights (what constitutes an Adaptation, what is non-commercial).

5. The differences with other similar open content licenses which have the same purpose, but use a different language and may become compatible with the BY-SA license. Efforts are indeed being led to reach compatibility by accepting that derivatives may be licensed not only under the same license but also under licenses which will have been recognized compatible. Four methods to improve compatibility between different open licenses and open licensed works are considered:
1. Cross-licensing and reciprocical compatibility clauses, with the example of the Free Art license,
2. Combination of works licensed under different licenses and partial compatibility between content with the example of the Digital Peer Publishing Licenses,
3. Dual-licensing and re-licensing or *de facto* compatibility between content by disappearance of one license, with the example of the Wikipedia migration from the GNU-GFDL to the CC BY SA 3.0 unported license,
4. Definition of common freedoms between licenses, one step backwards to go back to the basics.

After detailing external and internal legal incompatibilities and inconsistencies, the study evaluates their actual impact on contract formation and on the ability to make derivative works. Some consequences may be theoretical, minor or harmless, while some others may endanger the validity and the enforceability of the system, in some jurisdictions at least, and should be fixed. Before considering possible solutions to improve the system, it matters to assess whether correctives are really necessary, if there are a severe incompatibility and substantial cases where the licenses cannot be held valid and enforced. The impact could be that licensors may not be able to require their conditions to be enforced, and that licensees may not be able to claim the benefit from a grant which is more generous than copyright law, possibly spreading involuntary infringement.
It will be assessed which rights are at the entrance of the licensing process (when a Licensor licenses a Work) and at the exit (when a Licensee obtains that Work and wants to redistribute it or to make a derivative and become a Licensor). Licensors cannot license more rights than they own, and licensees cannot enjoy (and then further distribute or license) more rights than they were actually granted. As many other authors already noted, the proliferation of licenses and related information costs are jeopardizing free culture but also informed consent. Variations contained in future versions, in jurisdictions versions and in future versions of future compatible licenses cause legal insecurity because rights may not be the same for all the parties. Parties consent to one legal code, but cannot consent to all the other legal codes under which their modified work may be relicensed after the Share Alike compatibility clause, also because these differences are hidden in the licenses different versions, they are not accessible pieces of information.

Based on conclusions reached at various stages of this study, solutions mostly from logical and technical nature will be proposed to solve legal problems. Some elements could be drafted and implemented in the short-term without requiring too much effort. Other more substantial points could evolve in the long-term but require more research and development as well as consultation, particularly on the user interface, the definition of community guidelines, as well as for decisions involving changes in the substance of the provisions.
More technologies can be developed to better support the licenses requirements, such as attribution, the management of derivative works, the notice text, the definition of what constitutes the work which is being licensed, information on the licensor, etc.
I also propose options to improve the interface design. Following the model of the CC Public

Domain tools could solve problems of consent regarding consumer law requirements, limited representations of non-infringement and lack of identification of the contact person, being author or licensor.
The logic of the system could also better reflect positive freedoms and core clauses, before focusing on the options to be chosen to modify these freedoms. It could be considered to present first what is at the core of all licenses and will be modified by the choice of the licensor, instead of focusing on the options, qualitatively crucial, but quantitatively minor elements which may hide the core of the licenses. This change would be reflected in the license chooser and in the Commons deed.
Finally, I recommend reorganizing and redrafting the text of the licenses in order to rationalize and simplify the whole system. The text of the licenses should be shorter and in plain language. The Commons deed and the legal code could be combined in a single short and human-readable document presenting all the clauses in the form of clustered bullet points drafted in non-legal language illustrated by corresponding icons. But even before taking the important step to write only one short text, a reorganization of the legal code could improve the layout and the readability. It would be easy to reorganize and cluster thematics and to add subtitles. I also suggest changing the international porting process which introduces involuntary legal inconsistencies. Definitions could be drafted according to no legislation. Instead of being localized into jurisdictions, the CC porting process could take place within user communities and focus on translation and social governance by users rather than on legal normativity. Best practices could be defined and implemented within creative or user communities. A set of ethical principles described in an extended common deed or in a separated document may be more effective and accessible than a detailed doctrinal definition ported in a multiplicity of jurisdictions. Both judges and users could use these soft law guidelines to better understand and implement the licenses.

1. Introduction

The extension of copyright law duration and the expansion of its scope are reducing possibilities to access and reuse works, while digital technologies can make works more available instead of locking them even more[1]. Creative Commons aims at removing barriers to access and creativity by facilitating sharing of works[2]. To achieve this goal, Creative Commons provides standard licenses and other tools for authors to mark their works with the degree of freedom they wish to grant to the public, free of charge.

On the one hand, the movement born in 2002 has been relatively successful. More and more people have heard about Creative Commons[3], and millions of works, many of them created by famous artists and reputable institutions, or distributed on well-known websites[4], are available for free: permission has already been granted, icons makes it easy to identify these works and they are widely used by the "free culture" and "open access" movements.

On the other hand, the message and the strategy of the organization may lack of clarity and of a strong ideology to fix and redefine copyright[5]. Several licensing options are available, and the text of the licenses, what constitutes a "free" work,[6] or which rights are actually granted are not always well defined. Despite a user-friendly interface[7], this diversity of terms may have a chilling effect on the reuse of CC licensed works. The seven-year-old open content sharing system offers many different licenses to answer to the needs of various users community and the system is quite complex[8].

There are not only several options, but also several versions of the licenses, which are being

---

[1] See for instance Boyle James, "The Second Enclosure Movement and the Construction of the Public Domain", *Law and Contemporary Problems*, vol. 66, p. 33-75, 2003 and Lessig Lawrence, Free Culture - How Big Media Uses Technology and the Law to Lock Down Culture and Control Creativity, The Penguin Press, 2004, 348 p.

[2] Its motto on the website current homepage is " Share, Remix, Reuse — Legally. Creative Commons is a nonprofit organization that increases sharing and improves collaboration. " http://creativecommons.org/

[3] Among a population of 1115 first year students in the US surveyed for a research on Internet users skills in 2009, 7% of surveyed people had heard about Creative Commons, and the percentage is higher among those who share content on the Internet and especially among those who use sites like Flickr. Eszter Hargittai, Skill Matters: The Role of User Savvy in Different Levels of Online Engagement, Berkman Luncheon Series, Harvard Law School, 23-06-2009. http://cyber.law.harvard.edu/events/luncheon/2009/06/hargittai

[4] Gilberto Gil, MIT Open CourseWare, Al Jazeera, the White House, Flickr, Wikipedia...

[5] Elkin-Koren Niva, " Creative Commons: A Skeptical View of a Worthy Pursuit ", in Guibault Lucie & Hugenholtz P. Bernt (eds.), *The Future of the Public Domain*, Kluwer Law International, 2006 and " What Contracts Cannot Do: The Limits of Private Ordering in Facilitating A Creative Commons ", *Fordham Law Review*, vol. 74, p. 375, November 2005, see also Dusollier Séverine, " The Master's Tools v. The Master's House: Creative Commons v. Copyright ", *Columbia Journal of Law and the Arts*, 29:3, 2006, p. 271-293 and Chen Shun-ling, " To Surpass or to Conform – What are Public Licenses For?", University of Illinois Journal of Law, Technology & Policy, Vol. 2009, Issue 1, p. 107-139.

[6] Mako-Hill Benjamin, " Towards a Standard of Freedom: Creative Commons and the Free Software Movement ", July 2005. http://mako.cc/writing/toward_a_standard_of_freedom.html

[7] http://creativecommons.org/choose/

[8] Even if some licenses which were answering specific needs (Developing Nations, Sampling) have been retired.

translated into different languages and adapted to specific legislations[9]. It is unclear whether they contain exactly the same rights and restrictions and whether small language differences may have an impact on the rights actually granted and legal security of current users or the availability of works for future generations to access and build upon. The Share Alike provision is transmitted to derivative works, which can only be mixed among works licensed under the same or compatible conditions[10]. Besides, other provisions than the Share Alike clause including in non Share Alike licenses must be respected in derivatives. Therefore, not only these works are not compatible with works licensed under other copyleft licenses, but also possible problems may be transmitted to the future. Besides, other sources of legal uncertainty and incompatibility, as well as their actual or potential consequences, need to be evaluated, such as the enforceability of the licenses across jurisdictions with different and possibly inconsistent legislations, the variations between the licenses summaries and the actual text written in legalese language, the interoperability with other copyleft licenses.

The objective of this study is not to add to the critics and doubts of the skeptics of the system[11]without constructive propositions, but to make an objective evaluation of the licenses legal pitfalls and possible problems which may or may not arise, in order to make sure that works can be shared, accessed and reused with a maximum of certainty and security and a minimum of information and transaction costs. The marketing of a socially useful project must be supported not only by a clear political discourse, as suggested by critics of supporters of a strong public domain[12], but also by a solid legal infrastructure, which may require some adjustments to mitigate risks and improve legal certainty and compatibility in the future.

This research aims at identifying legal issues and assessing the actual consequences of inconsistencies of a system submitted to multiple constraints: users community requirements, national legislations diversity, international private law complexity, differences between a multiplicity of licenses. When possible and useful, this research will try to propose solutions to legal pitfalls and incompatibilities in order to maintain the original goals of legal security and simplicity of the open licensing framework. Indeed, "the establishment of a reliable semi-commons of creative material that can be used by others without worrying about the overly restrictive and complicated law of copyright (…) is central to the goal of Creative Commons"[13].

---

[9] See the Creative Commons international "porting" process description.
http://creativecommons.org/international/

[10] Here is the definition of Share Alike, in the human human-readable summary of the Legal Code, and in the Legal Code (the full license):
"If you alter, transform, or build upon this work, you may distribute the resulting work only under the same, similar or a compatible license." http://creativecommons.org/licenses/by/3.0/
"You may Distribute or Publicly Perform an Adaptation only under the terms of: (i) this License; (ii) a later version of this License with the same License Elements as this License; (iii) a Creative Commons jurisdiction license (either this or a later license version) that contains the same License Elements as this License (e.g., Attribution-ShareAlike 3.0 US); (iv) a Creative Commons Compatible License. If you license the Adaptation under one of the licenses mentioned in (iv), you must comply with the terms of that license." http://creativecommons.org/licenses/by-sa/3.0/legalcode

[11] See Koelman Kamiel, "Waarom Creative Commons niet kan werken", *Computerrecht* 2009, p. 112 and Farchy Joëlle, "Are free licences suitable for cultural works?", *European Intellectual Property Review*, 2009, vol. 31, n° 5, pp. 255-263.

[12] See *op cit* Chen, Dusollier, Elkin-Koren, Mako-Hill.

[13] Pallas Loren Lydia, " Building a Reliable Semicommons of Creative Works: Enforcement of Creative Commons Licenses and Limited Abandonment of Copyright ", *George Mason Law Review*, vol. 14, p. 271, 2007.

1.1 Sources of legal incompatibilities

Creative Commons licenses have been designed to facilitate the use and reuse of creative works by granting some permissions in advance. However, the system is complex, with a multiplicity of options, formats and versions, making it difficult to understand which subject matter and rights are exactly covered. There is a risk to see resources intended to be part of an intellectual commons pool underused and transaction and information costs increased, while the initial goal of the framework was to provide simple tools, support legal security and foster sharing, reuse, access and creativity.

The risk of license proliferation, or of not being able to remix works licensed under close, however different open content licenses requiring derivatives to be licensed under the same license, has been identified by many scholars and users, including the founder of the movement[14]. It is inherent to the copyleft provision and cross-licensing policies may solve the issue and avoid open content ghettoisation. Even not all works available under one of the Creative Commons licenses can be combined without further negotiation because not all licenses options are compatible among each other: "an unsolvable dilemma"[15]. The multiplicity of Creative Commons licensing options increases confusion and information costs besides leading to frustrating internal incompatibilities[16]. Can the proliferation of licenses lead to the anticommons[17], with fragmented, underused resources which cannot be recombined?

Besides these visible sources of incompatibility between works, there are also differences within each license which might be sources of inconsistencies but which are not visible to the average user, first among the various formats, and second among the local adaptations.

The human-readable summary, which is visible and easily readable accessible, but not legally binding, does not contain the same level of details than the legal code, which is much longer and more detailed. Provisions from the core grant do not appear in the title of the licenses which only display the optional provisions. Are users aware of the conditions to which they really consent? What are the risks for the licenses validity and could the infrastructure be improved to increase awareness and informed consent without losing the simplicity of the two-tier system?

---

[14] "The project of private ordering a commons, however, faces a number of significant challenges. Perhaps the most important is to assure that freely-licensed creative work can, in a sense 'interoperate'. If work licensed under one free public licence cannot be integrated with work licensed under a second free public licence, then a significant part of the potential for free licensing will be lost.", Lessig Lawrence, "Recrafting a Public Domain" (2006) 18 *Yale Journal of Law & Humanities* 56, p. 77.

[15] Séverine Dusollier, Sharing Access, *op cit*, p. 1425 et s.

[16] Niva Elkin-Koren, What contracts can't do, *op cit*, p. 51 et s. and Shaffer Van Houweling Molly, "Cultural Environmentalism and the Constructed Commons," *Law and Contemporary Problems* 70, Spring 2007, p. 23-50.

[17] The tragedy of the anticommons has been coined by Michael Heller: when too many owners hold a right of exclusion, rights clearance is too difficult or even impossible, all the more for products and collective works which require to assemble many preexisting works: Heller Michael, "The Tragedy of the Anticommons: Property in the Transition from Marx to Markets", *Harvard Law Review*, Volume 111 (3), January 1998, p. 621-688; Heller Michael, *The Gridlock Economy: How Too Much Ownership Wrecks Markets, Stops Innovation, and Costs Lives*, Basic Books, New York, 2008, 304 p.

Differences between the various licenses, especially between adaptations to jurisdictions' legislations, are not very accessible to the public, and their impact has not been studied yet. In order to be compatible with the legislation of each jurisdiction, their terms are adapted and are thus all slightly different; then, how can they be declared compatible among each other? Is the Creative Commons porting process generating additional difficulties, or are inconsistencies a necessary harm due to the fact that copyright is a national matter, but which do not worsen a cross-national differences situation which can't be solved by private regulation but only by public ordering? What happens if users are not aware of differences between the licenses? Is there a risk of breach of contract in addition to copyright infringement?

This study will be presenting the different licenses (chapter 2), identifying the various possible sources of legal incompatibility (chapter 3), assessing their actual impact (chapter 4) and finally proposing and evaluating options to mitigate risks and improve compatibility, consistency, clarity and security (chapter 5). Indeed, the goal of the study is not to criticize the project[18], but to identify potential problems and attempt at solving them before they become acute.

Are these incompatibilities and possible sources of inconsistencies a real threat to the security and the sustainability of the system? Could the Creative Commons system be simplified and what could be the possible solutions to improve rights clearance, licensing information and legal security for licensors and licensees? Could sectorial user communities play a role in a possible reform or taylorisation of the Creative Commons system and how? What are the best ways to deal with licenses incompatibility and proliferation problems which are also happening in the free and open source software communities [19]? Would the definition of common principles and guidelines to govern the licenses solve legal problems?

1.2 Scope, methodology and outline

In order to compare the licenses and assess the impact of their differences, we choose as a starting point the legal deed of the licenses version 3.0 unported, which will be considered as the standard to be compared with the other formats, versions and jurisdictions.
The unported license is the text that jurisdictions are translating and porting to their local law[20]. They are expected to vary as little as possible from this standard, in order to stay as compatible as possible. Variations are justified only to the extend that they are required to ensure local validity[21]. Since version 3.0, the unported license refers to concepts as defined in

[18] Following other critics of the strategy of the movement identifying "potential defects and risks of the model (...), it helps to counteract possible criticisms that might undermine the very objective of the action", in Dusollier Séverine, The Master's Tools v. The Master's House: Creative Commons v. Copyright, *Columbia Journal of Law and the Arts*, 29:3, 2006, p. 273.
[19] The Open Source Initiative drafted a report on license proliferation (http://www.opensource.org/proliferation-report) and approves some licenses as "open source": http://opensource.openmirrors.org/node/365.html
[20] On the Creative Commons international (CCi) porting process, see Maracke Catharina, "Creative Commons International. The International License Porting Project – origins, experiences, and challenges", in Danièle Bourcier, Pompeu Casanovas, Mélanie Dulong de Rosnay, Catharina Maracke (eds.), *Intelligent Multimedia. Sharing Creative Works in a Digital World*, Series in Legal Information and Communication Technologies, Vol. 8, European Press Academic Publishing, Florence, 2010, p. 67-88.
[21] "For compatibility purposes, you may not modify the license beyond what is necessary to accomplish compliance with local law." http://wiki.creativecommons.org/Legal_Project_Lead_produces_a_first_draft

international treaties. Before version 3.0, the unported license was called generic, and it was based on US copyright law definitions.
The comparison between licenses will be systematic and highlight all the differences between formats and versions, while it will focus only on key provisions of the core grant to illustrate the differences between jurisdictions versions and with other open content licenses. Differences will be analyzed among the successive unported and jurisdictions versions of the Creative Commons core licensing suite combining the following optional elements: Attribution (BY), Non Commercial (NC), Non Derivative (ND) and Share Alike (SA). The Sampling licenses, the Developing Nations license, the Founders Copyright, the Public Domain Dedication, the CC0 and the CC+ protocol will be analyzed to the extend their characteristics can be useful for the purpose of the study, without leading a systematic comparison in order to identify differences or incompatibilities.

This legal study on the Creative Commons licensing system pitfalls, risks and potential incompatibilities starts by a presentation of the CC movement and the licenses (section 2.1) which are made available from a online license chooser in a multiplicity of formats (section 2.2.1) and options (section 2.2.2) flavoring core clauses (section 2.2.3). We will then analyze their legal nature and effects (section 2.3).

After a description of the licenses diversity from the viewpoint of the user downloading a license from the interface (chapter 2), the study will detail the identified and potential sources of incompatibilities between all the licenses which are actually available (chapter 3), from the identified sources which are easy to grasp and manage, to the less visible and more problematic differences:

- The differences between the languages contained in the various formats of the licenses (section 3.1),
- The evolution between the four successive versions, when clauses have been added or removed for improvement and rationalization purposes (section 3.2),
- The variety of options, preventing to combine two works licensed under different license optional elements and causing a fragmentation in the commons pool and philosophy (section 3.3),
- The opportunities and caveats offered by the porting process of the unported licenses, which legal deed has been adapted into the language and legislation of over 50 jurisdictions (section 3.4),
- The differences with other similar open content licenses, in the light of the work achieved of the Open Source Initiative[22] on ongoing or possible negotiation process toward compatibility with the Creative Commons Attribution Share Alike license: the GNU Free Documentation License (GFDL)[23], the copyleft Free Art License (FAL)[24] and the Digital Peer Publishing Licenses (DPPL)[25]. Four possible solutions to the problem of license proliferation will be analyzed:
- Dual-licensing and re-licensing with the example of the Wikipedia migration process,
- Cross-licensing provisions,

[22] Report of license proliferation, *op cit.*
[23] GNU Free Documentation License. Version 1.3, 3 November 2008. http://www.gnu.org/copyleft/fdl.html
[24] Free Art License 1.3 (FAL 1.3), 2007. http://artlibre.org/licence/lal/en
[25] Digital Peer Publishing Licence (DPPL), Version 3.0 – November 2008. http://www.dipp.nrw.de/lizenzen/dppl/dppl/DPPL_v3_en_11-2008.html

- Combination of content licensed under non-compatible terms and finally
- The definition of standard or “essential freedoms” to categorize open content licenses, with a proposal in the light of the initiative of the Definition for Free Cultural Works and Licenses[26].

The impact of pitfalls and incompatibilities will then be analyzed with a spotlight on the consequences for the licenses validity and enforcement for creators, users and intermediaries’ legal security, as well as for the ecosystem simplicity and balance.
The legal validity of the agreement will be analyzed from the viewpoint of licensors and licensees with a description of contract formation and how this framework applies to the ability to consent to Creative Commons agreements (section 4.1).
Licensees and intermediaries’ legal security as well as the ability to actually use all works and make derivatives will then be evaluated (section 4.2), focusing on a selection of clauses of the core grant which differ between versions and jurisdictions: moral rights, database rights, warranties and collecting societies.

The concluding chapter of the study will consider and assess several possible solutions to correct pitfalls and incompatibilities, mitigate or limit consequences and try to simplify the system, including improving the interface design and the language of the licenses and relying on technology and coordination by intermediaries.

[26] http://freedomdefined.org/Definition

2. Creative Commons licenses diversity

The expression “Creative Commons” designates an organization, a set of copyright licenses and a trademark. The set of Creative Commons licenses proposed to the public by the Creative Commons organization are private agreements which apply on the top of the law as a form of exploitation of rights emerging from copyright. The Creative Commons organization is promoting Creative Commons licenses, aiming at supporting the needs of various communities who want to share and reuse works more easily and under more permissive terms than allowed by default copyright law. The licenses are free, and come with a set of tools, logos, educative material and machine-readable code.

We will describe Creative Commons infrastructure (section 2.1.1) and policy (section 2.1.2). The licenses are made available to the public in different formats (section 2.2.1) and combination of optional elements (section 2.2.2) around core clauses (section 2.3). Beyond the core clauses constituting the common denominator of the licenses, some provisions are optional and lead to a puzzle of optional elements (section 2.2.2) which are to be selected from the license chooser interface and are combined around the main clauses (described under section 2.2.3). The assemblage of the optional elements around the core clauses producing one of the currently available six licenses. Licensors may or may not request their work to be used for non-commercial purposes only, they may or may not request their works to be used in a non-derivative way only, and may or may not request the derivatives to be licensed under the same conditions. Based on the answers to these questions, the currently six licenses[27] are combining none, one or two of the three optional elements Non Commercial, No Derivative Works and Share Alike:

- Attribution (BY)
- Attribution - Share Alike (BY SA),
- Attribution - No Derivative Works (BY ND),
- Attribution - Non Commercial - No Derivative Works (BY NC ND),
- Attribution - Non Commercial (BY NC),
- Attribution - Non Commercial - Share Alike (BY NC SA).

Several *incremental* versions have been made available, in order to rationalize the licenses text. Some of the clauses have been deleted or added between the four versions, namely versions 1.0, 2.0, 2.5 and 3.0. The licenses are being first released by the organization in generic or *unported* versions: first based on US law definitions, then based on international conventions definitions. Finally, *jurisdictions’* versions of the licenses are being made available: the organization uses the term of ‘legal porting’ to convey the idea that clauses of the *unported* version are translated and localized to improve compatibility with local languages and national legislations after a legal adaptation. We will study these questions in section 3.3 (incremental versions from 1.0 to 3.0, named thereafter “versions”) and section

[27] Version 1.0 of the licenses had one additional optional element, Attribution, which ceased to be optional to become a part of the core grant from version 2.0 (more details in section 3.2), thus reducing the number of available licenses from eleven to six.

3.4. (localized versions of the unported version, porting the licenses legal code to the legislation of over 50 jurisdictions, named “jurisdiction licenses” or “ported licenses”).

However, for the methodological purpose of this study, we will start by considering in this chapter 2 only the differences which are immediately visible from the working of the system; when using the license chooser interface, a license is generated in various layers or formats (2.2.1) according the optional elements (2.2.2) which have been selected to modulate the core clauses (2.2.3) of the license available in its current available version, namely version 3.0.

When not mentioned otherwise, and in order to define a standard or median point of comparison with the other licenses of the system to be studied, we will analyze the CC BY NC SA 3.0 unported license. Indeed, this license, in its unported version as released by the headquarters intending to reflect international texts such as the Berne Convention, contains almost all the existing clauses[28] after the previous incremental versions and before the localized versions, the jurisdictions’ licenses.

It is important to differentiate the *median license* containing the language of *all the clauses*, from the core, basic or minimum freedoms offered by *all the licenses*. This notion was not obviously displayed in the early years of the organization, when it did not have a clear policy (section 2.1.2). It has been defined as the right to share the work for non-commercial purposes only, with attribution and without modification (BY NC ND), which can be augmented by more freedoms by replacing ND with SA or by removing NC and/or ND optional elements.

After a review of the licenses infrastructure and policy (section 2.1), we will describe the licenses as generated by the system in different formats (section 2.2.1), with optional provisions (section 2.2.2) wrapped around main clauses (section 2.2.3). Once we will have a clearer picture of the object of our analysis, the licenses, we will be analyzing and interpreting their legal nature (section 2.3) in a systemic way. Indeed, the licenses are used by agents and circulate along with works. They are supporting a complex system, the pool of works made available to the public for sharing and reuse, which this study tries to keep sustainable, to allow agents to distribute and reuse works with the lowest costs and risks possible. The legal nature of the licenses should be qualified according to contract law in order to evaluate how they apply and what can be their effects between the parties involved, licensors, licensees, authors, the public, potential future users. It should be qualified who has a relation with whom, what kind of relation it is, casual or contractual, permissive or with duties (section 2.3.1), and how this relation impacts subsequent partners and offspring in case of derivative works. Indeed, because of the viral nature of the contracts[29] and of the copyleft Share Alike provision (section 2.3.2) binding subsequent users, works released under a CC license continue to carry the licenses freedoms and obligations.

Describing the licenses (section 2.1) as well as identifying their various features (section 2.2) and how they function legally (section 2.3) will allow to describe the sources of potential incompatibility in chapter 3.

---

[28] With the exception of the compatible licenses clause which is available only in the BY-SA 3.0.

[29] Viral contracts are following their product, they have been described by Radin Margaret Jane, “ Humans, Computers, and Binding Commitment ”, *Indiana Law Journal*, vol. 75, p. 1125-1161, 2000.

2.1 Creative Commons: an organization and a set of licenses

Creative Commons is a non-profit organization which has been created in the United States in 2001, and provides since 2002 free copyright licenses for authors to mark their works with the degree of freedom they wish to grant to users.

In this section, we will present the licenses infrastructure and tools (section 2.1.1), and how Creative Commons policy is being defined, oscillating between standardization and diversity (section 2.1.2). While lacking the flexibility and the personalization of tailored-made items, standardization has numerous advantages: it lowers information and transaction costs and fosters interoperability between industrial products. This general statement related to technical standardization is also applicable to the Creative Commons licenses and organization, which provides ready-to-use tools. This section will assess if Creative Commons is a standard on the technical, legal and policy levels. Indeed, standardization aims at creating interoperable products and in order to work properly, the licenses framework needs to interoperate nicely, both internally among the various layers and versions, and externally with the legal system.

2.1.1 The licensing infrastructure: a technical standard

The licenses were launched in December 2002 and every year or almost, a new product or a new version is being released, in the same vein than software with upgrades, to correct bugs or address niches. Like a technical standard, the CC system contains several complementary elements: a user interface or license generator, a multiplicity of licenses and tools to identify and remix licensed works, machine-readable code, specifications such as FAQs and educational material explaining how to use the licenses and marketing products in the form of short movies and comics explaining why to use the licenses.

The initial version 1.0 was offering eleven licenses, which have been reduced to six licenses after the revision leading to version 2.0 making the Attribution element non-optional and part of the core grant. Versions 2.5 and current version 3.0., the only one available from the license chooser interface, did not modify the number of licenses but only the core clauses. More licenses outside the core suite of 11 and then 6 licenses have been made available (the Sampling and the Developing Nations licenses) and then withdrawn because they were not granting the common freedom to share non-commercially[30]. Finally, the Public Domain Dedication based on US law has not been formally retired, but has been replaced by the CC0 waiver, another tool, this time aiming at placing works as close as possible to the public domain and thus not based on US law only.

Unlike tailored copyright licenses written by lawyers for specific and unique needs comparable to "*haute couture*", Creative Commons provides six "*prêt-à-porter*" or "ready-to-

[30] Retired licenses are listed at http://creativecommons.org/retiredlicenses. This page explains that all licenses "guarantee at least the freedom to share non-commercially". More detailed explanation on the fact that these licenses were not granting core freedoms or "minimum standards" of the open access movement: http://creativecommons.org/weblog/entry/7520 and *infra* in sections 2.1.2 and 3.5.3.

wear" texts aiming at answering most needs while minimizing the number of available "sizes" or "colors". Indeed, the licenses are a patchwork of eight core clauses, with variations among the additional clauses corresponding to the available options selected through the generator which produced a license in various formats (section 2.2.1). These options will be described in the next section (section 2.2.2). Their assembly constitutes the name of each of the licenses. These options flavor a sauce base, a core grant which is not expressed in the title of each of the licenses: the non-exclusive right to reproduce, perform and distribute the unmodified work for non commercial purposes. The clauses of this core grant will be studied in more details (section 2.2.3).

The Creative Commons model intends to be simple and easy-to-use. But there are actually not only six combinations of options, even when addressing only the current core unported version, disregarding previous versions and licenses outside the core system.
The core licenses are the 11 and then 6 licenses, without including the other tools proposed by the organization, such as CC0 or the Sampling licenses.
The six core unported licenses have been or are to be translated and adapted to over 50 jurisdictions. Previous versions of the licenses continue to be in use. As explained earlier, the unported licenses are the standard version based on international conventions definitions before the localization porting process leading to jurisdictions versions which will be studied in detail in section 3.2, as sources of potential incompatibilities and inconsistencies.
The purpose of having jurisdiction-specific licenses is to provide a linguistic and legal translation, as well as to increase access, acceptability and understanding by users and judges who need to interpret the licenses in local jurisdictions. The internationalization or porting process also provides local teams of project leads. Beyond ensuring the translation and legal porting of the legal code, jurisdictions project leads work with local user communities and governments to explain and promote the licenses. Jurisdictions teams also collaborate with CC headquarters staff to perform research, provide suggestions to improve the licenses' clauses and overall infrastructure, report on questions, use cases and issues arising in their jurisdiction, translate and create educational material and constitute a network advising on questions affecting user communities around the world.

Several applications have been developed to support[31] the legal tool in the networks (search services[32], a rights expression language[33], a remix website[34]) and the license terms are

[31] This intrication between code and law reflects the scholarship of Creative Commons' founder. See Lessig Lawrence, *Code and other laws of cyberspace*, Basic Books, New York, 1999, 297 p.
[32] The goal of having a machine-readable format of the licenses is to have a proof of concept of the semantic web and allow users to search for works according to their licensing conditions, so that they can be reused and integrated: use for commercial purposes or not, modify or not. The initial project ccNutch was a search engine based on Nutch open source technology and RDF, indexing only results with CC metadata and displaying works according to their license elements (see press releases http://creativecommons.org/weblog/entry/4028 and http://creativecommons.org/weblog/entry/4388). The technology has been integrated as a plugin of the Firefox browser (see the 2004 press release at http://creativecommons.org/press-releases/entry/5064 and more explanations at http://wiki.creativecommons.org/Firefox_and_CC_Search). Now, ccSearch at http://search.creativecommons.org/ is a portal which aggregates results provided by CC enabled search engines provided by Google and Yahoo! for web results, Flickr and Wikimedia Commons for images and Jamendo for music, among other databases and repositories.
[33] A Rights Expression Language is a abstract model containing the syntax and the semantic needed to describe copyright permissions and authorizations and build automatized applications, such as the search engines described just above, or Digital Rights Management systems. RDF is the standard to express semantic information on the web. ccREL uses RDFa to express semantic information about objects' licenses. For more information, see http://wiki.creativecommons.org/RDFa and http://wiki.creativecommons.org/ccrel, the W3C

embedded in machine-readable code or metadata. The licenses are declined into four layers or formats (section 2.2.1):

- A button to be displayed on works' websites and physical supports, containing a link to the license human-readable summary, the commons deed,
- Machine-readable code embedded in the HTML specifying the logo and available from the deed, containing metadata to be processed by search engines to locate works according to their licensing conditions,
- A human-readable summary of the license's core freedoms and optional restrictions, accessible from a link inside the logo,
- The legal code, e.g. the full license, accessible from a link at the bottom of the human-readable summary.

Due to the success of the licenses which are applied to more than 250 million objects on the Internet as of June 2009[35], the Creative Commons licenses are becoming a *de facto* standard of open content licensing[36] and more broadly for collaboration on the Internet[37]. Creative Commons as an organization is contributing to the technical standardization of the web[38]. The licenses could become *de jure* standards: governments releasing public sector information under one of the Creative Commons licenses may be mandating or recommending the use of the licenses for works they create or subsidize.

Creative Commons organization and licenses intend to cover the public domain and the "no rights reserved" perspective, and some of the spectrum of rights between that and the "all rights reserved" approach, with a set of standard licenses combining various options and containing less restrictions than the full spectrum of copyright protection applicable by default to every work as soon as it is created, thus: "some right reserved".

---

specification submission http://www.w3.org/Submission/ccREL/ and the article by Abelson Hal, Adida Ben, Linksvayer Mike, Yergler Nathan, ccREL: The Creative Commons Rights Expression Language, Communia First Workshop, Torino, January 2008. http://www.communia-project.eu/node/79

[34] ccMixter at http://ccmixter.org/ is "a community music site featuring remixes licensed under Creative Commons where you can listen to, sample, mash-up, or interact with music", providing a useful "Derivation History" for each track, a "Remix History Chart" of samples used, which could be applied to other domains than music to trace pre-existing contributions and derivative works.

[35] For information about adoption metrics and statistics, see http://wiki.creativecommons.org/Metrics and http://wiki.creativecommons.org/License_statistics.

[36] For instance the "recognition of Creative Commons as the standard for sharing" in the Google Book Settlement: Linksvayer Mike, "CC and the Google Book Settlement", *CC blog*, 16-11-2009. http://creativecommons.org/weblog/entry/19210

[37] The "TCP/IP of collaboration and content layer" for CC CEO Ito Joi, "Creative Commons: Enabling the next level of innovation", *What Matters*, McKinsey & Co, 30-10-2009. http://whatmatters.mckinseydigital.com/internet/creative-commons-enabling-the-next-level-of-innovation. Original unedited version: http://joi.ito.com/weblog/2009/10/30/innovation-in-o.html

[38] See *infra* footnote 30 about RDFa.

2.1.2 Creative Commons policy strategy: not quite a legal standard

We just saw how Creative Commons can be defined as a standardized infrastructure providing a set of tools to distribute, access and reuse free works and develop the commons. We will now briefly describe political and legal implications of the choices at the origin of the available options, and critiques resulting from these choices coming from the free software community. It is interesting to compare the strategic choices of Creative Commons with the open source and free software community at various levels. First, CC claims to follow the model of its predecessors for non-software content. Second, the movement is very successful in federating communities and adopting a single standard of freedom.

The policy message of Creative Commons is that to provide an alternative to full copyright. But because so many licenses are available without defining a core freedom specifically enough[39], Creative Commons has been accused on the one hand of lacking of a core message and on the other hand of not being free enough. Indeed, many scholars of the public domain and actors of the free and open source software communities have expressed critical views of Creative Commons tools and movement[40]. We will use here only the subset of these critiques which is relevant to the diversity/standardization dichotomy and will highlight future developments on licensing options[41]. Indeed, the high number of options, coupled with an absence of a clear definition of the core freedoms of a CC license, are sources of incompatibilities and ideological critiques may provide useful hints to improve the system and solve some incompatibilities issues by making the system a true legal standard.

For Niva Elkin-Koren, "The legal strategy (…) facilitates a far-reaching coalition among libertarians and anarchists, anti-market activists and free-market advocates. At the same time, however, Creative Commons lacks of a (…) clear definition of the prerequisites for open access to creative works. The end result is ideological fuzziness."[42] The diversity of licensing options still increases information and transaction costs. As the goal of CC is to minimize information and transaction costs, the licenses could benefit from more standardization: "Creative Commons' strategy presupposes that minimizing external information costs is critical for enhancing access to creative works. It seeks to reduce these costs by offering a licensing platform. Yet, the lack of standardization in the licenses supported by this licensing scheme further increases the cost of determining the duties and privileges related to any specific work. This could add force to the chilling effect of copyrights." She regrets the "lack of a clear definition of the *commons*" [43].

A lot of energy was involved to reach consensus and a shared definition of free software, in order to offer only one option (corresponding to Attribution Share Alike), but the FLOSS movement includes many different clauses and also permissive licenses, roughly corresponding to CC BY, and to the Public Domain. CC did not choose to offer only one license. Providing only one CC license would:

---

[39] See Mako-Hill, *op cit*.
[40] For a review of existing criticisms, see Chen, *op cit*.,
[41] See *supra* section 3.3 on the incompatibilities between options and section 3.5.3 on the definition of freedom.
[42] Elkin-Koren Niva, What contracts can't do, *op cit*, p. 6.
[43] Elkin-Koren Niva, A Worthy Pursuit, *op cit*, p. 10.

- Satisfy one clear definition of freedom,
- Avoid one of the risks of incompatibility, the incompatibility between works licensed under different options,
- Provide guidance to users, instead of re-creating high information costs or barriers to entrance when it is the first time to select a license or use a licensed work.

On the contrary, the organization chose to offer the choice between various levels of freedoms[44] to attract different audiences to free culture, including authors who are not ready to give away commercial and derivative rights, but are willing to otherwise share their works with the public. The strategy to satisfy various needs and communities and the related ideological fuzziness cause incompatibilities because too many options are available. Also, a clearer definition of what constitutes freedom could reduce information costs and legal uncertainty if users do not fully realize to what combination of options they consent. Indeed, if there was a strong conceptual definition of what principles constitute freedom, and few variations from that core, there will be fewer incompatibilities.

Still, it might be difficult to reach a consensus on what constitutes freedom (and thus define a core message and strategy) between users who have multiple roles and diverse expectations. It took a long time to CC to set standards, as recalls Shun-ling Chen[45]. First, CC recognized the CC standard, the freedom to share works non-commercially[46], by withdrawing the licenses which were not ensuring this minimum grant[47]. Second, CC recognized a higher standard of freedom by clearly identifying which of its licenses comply to this standard with a new button, "approved for free cultural works"[48]. For Shun-ling Chen assessing the differences between the legal strategies of the Free Software Movement and Creative Commons more flexible model, Creative Commons is more about the freedom of authors than the freedom of a user community[49]. Of course actors of the movement and members of the public at large are both creating and consuming content and the distinction between authors and audience is not as sharp as in the analogue age. But a shift from trying to fulfill the wishes of the authors to giving more importance to the needs of the users might rationalize the system and reduce the number of options, sources of incompatibilities, and make it more secure for users.

We will detail the available licenses in the next section 2.2, and will come back to this notion of standard of freedom in the next chapter when we will be analyzing options for the compatibility with other open content licenses (section 3.5.3). Options rationalization and a more users-oriented approach will be part of the solutions proposed in the final chapter 5 of the study.

---

[44] Copyleft Attitude community at the origin of the Free Art License opposes their "choice of freedom" (only one license offering a core freedom) to Creative Commons' "freedom of choice" (several licenses offering several degrees of freedom). See Vodjdani Isabelle, Le choix du Libre dans le supermarché du libre choix, 2004, 2007. http://www.transactiv-exe.org/article.php3?id_article=95

[45] *Ibidem* p. 126-130.

[46] Lessig Lawrence, " CC in Review: Lawrence Lessig on Important Freedoms ", Lawrence Lessig, *CC News*, December 7th, 2005. http://creativecommons.org/weblog/entry/5719

[47] Lessig Lawrence, " Retiring standalone DevNations and one Sampling license ", *CC News*, June 4th, 2007. http://creativecommons.org/weblog/entry/7520

[48] Linksvayer Mike, " Approved for Free Cultural Works ", *CC News*, February 20th, 2008. http://creativecommons.org/weblog/entry/8051

[49] Chen Shun-ling, *op cit*, p. 121.

2.2 The different licenses available

This section describes the license system formats as well as its optional and non-optional clauses. Behind the optional elements, a core set of permissions allows verbatim non-commercial sharing. This core grant is not recognized as free as in free software and free culture because the freedom to make changes is not granted. Only two out of the six Creative Commons licenses (CC BY and CC BY SA, as well as CC0) are recognized as "free culture licenses"[50] because they grant the freedom to distribute derivative works, with or without permissible restrictions such as copyleft (Share Alike in Creative Commons vocabulary), the transmission of licensing conditions from original works to their derivatives.

The licenses are made available from the license chooser interface in four different layers or formats (section 2.2.1): a button, HTML code, a summary and a longer text, the actual license. After describing these formats, we will present the different options or license elements (section 2.2.2) which complement the core clauses (section 2.2.3). Thus, we will have a complete overlook of the various unported licenses which will allow further comparison with the other instances of the licenses to detect differences and incompatibilities among formats and options, the visibly different licenses.

2.2.1 The licenses formats

According to the CC website[51], "Creative Commons licenses are expressed in three different layers or formats: the Commons Deed (human-readable code), the Legal Code (lawyer-readable code); and the metadata (machine readable code)."

A forth item can be added to the list, the Notice Button, the first format generated by the system linking to the other ones. It is often the first instantiation of the license visible by the public, both the licensor choosing a license and the potential licensee seeing the button next to a work she might want to reuse. By answering the questions on the license selection interface to combine optional elements, prospective licensors obtain a link to the license corresponding to their choice. They are prompted to attach this license to their works to indicate which rights they grant to the public and which rights they reserve, by inserting on their website some HTML code which is being delivered by the license selection interface[52]. This piece of code represents a button with the Creative Commons logo and the icons corresponding to the options selected. The image contains a link to the license which has been selected, for instance:

Each of the 6 combinations forming a CC license is available in four formats linking to one

[50] http://freedomdefined.org/Definition
[51] http://wiki.creativecommons.org/FAQ
[52] For instance:
This work is licensed under a Creative Commons Attribution 3.0 Unported License.

other:

- A notice with a button displaying icons of selected optional elements,
- The machine-readable code,
- The human readable code with icons,
- The legal code.

The notice button can be the only format which is directly visible to the end-user visiting a website or looking at the printed copy of a work[53]. It is a major asset of the organization, displaying its logo and trademark, and acting as a symbol, a signal that the content can be shared and reused for free. Specific conditions are just a click away as the notice button contains a link to the license.

It should be noted that the initial version of the button was the same for all the combinations, only the CC logo which HTML is embedding a link to the human readable code. Critiques on the lack of visibility of a core message hiding the options contributed to the re-design of the button, this time integrating inside the CC logo either one, or two or three icons representing options of each license. A source of confusion was, and still is despite the displaying of the options icons in the button, that many users do not distinguish among the options and simply consider that a work is available under a, if not "the", CC license, without indicating which one. However, the code delivered by the interface contain not only the logo but also a sentence indicating for instance "This work is licensed under a Creative Commons Attribution 3.0 Unported License", the notice text. Specific design efforts should continue to be led to clarify what license is applied for all users even less mindful ones.

A source of misinformation and confusion is the lack, on many websites, of a proper notice next to the button. We can deduce from this absence that despite CC tutorials and FAQs, some authors or web designers either copy the button from other websites without using the interface to select their option and generate their code, or delete the sentence. Pallas-Loren[54] uses the term "notice" to refer to the combination of the button and the sentence accompanying it to stipulate that the work is available under a given license. We use the expression of "notice button" to designate the first format under which the licenses are being made visible to the public, both as licensor getting a piece of code from the interface and potential licensee seeing a logo and a sentence. This first format comes in addition to the three formats usually identified (human-readable, machine-readable and lawyer readable). It is very important as it may be the only format that a licensee will pay attention to, a button with icons and a sentence generated by the interface: "This work is licensed under a Creative Commons Attribution 3.0 Unported License".

We will now discuss the importance of one word in the notice sentence, the word *work*. Indeed, the license is applied to a specific work. And the text generated by the interface containing the notice sentence and the HTML code of the button should be copied next to a work to indicate its licensing conditions: "Copy the text below to your Web site to let your visitors know what license applies to your works", says the CC website above the text to be pasted to insert the notice button on a website. Thus, the clarification of what exactly is this *work* by the licensor when pasting the code on her website is a considerable and underestimated matter. Otherwise, it might not be clear what work is licensed. Is the "work"

[53] A text notice only may be present in place of the notice button.
[54] Pallas-Loren, *op cit*, p.12.

the website as a whole, some of the individual works placed on the website, for instance only the text but not the images? Most users do not specify what works on their website are covered by the license they chose, even if they use the sentence in their notice and not only the button. This lack of specification of what is actually covered may impact the validity of the agreement (section 4.2). A convenient and broad way to specify what is intended to be covered is to use the same sentence as on the CC website: "Except where otherwise noted, content on this site is licensed under a Creative Commons Attribution 3.0 License". This is not the sentence which is currently generated by the interface, but this could be changed and several options (single work, general website) HTML easy to copy/paste could be offered.

The name of the license within the notice sentence and the notice button itself contain a link to the human-readable code of the license. For instance, “Creative Commons Attribution 3.0 License” will link to the Commons Deed at http://creativecommons.org/licenses/by/3.0/. This link to the license human-readable format is the central element of all the formats. When correctly placed next to an identified work, users will be able to read under which conditions the work has been made available to the public by its licensor.

The Commons Deed or human-readable code contains a summary of the license main provisions: the options and some of the core clauses. CC FAQs describe the Commons Deed as “a summary of the key terms of the actual license (which is the Legal Code)—basically, what others can and cannot do with the work. Think of it as the user-friendly interface to the Legal Code beneath. This Deed itself has no legal value, and its contents do not appear in the actual license.” It is interesting to note the connotation of the chosen name of Commons Deed: “a deed is commonly understood to be a permanent conveyance of an interest in land”[55].

The Commons Deed is available in around 50 languages which are prominently listed at the top of the webpage. Linguistic diversity is being taken seriously into account by CC who offers for instance several Chinese, English, Spanish and French translations as these languages are spoken in different jurisdictions. However, if any user accessing to a Commons Deed in a foreign language can easily access to a translation in her mother language by clicking on the link at the top of the page, the first version which will be displayed will be the one of the jurisdiction chosen by the licensor and the licensee may read a translation presenting differences. As explained further for the differences between Legal Codes jurisdictions versions (section 3.4), the scope of rights granted by the Licensor in one jurisdiction may well not perfectly match the scope of rights granted to the Licensee reading another jurisdiction’s version. For instance, the Canadian English version allows “to copy, distribute and transmit the work” while the other English versions allow “to copy, distribute, display, and perform the work” and the Italian version to “comunicare al pubblico, esporre in pubblico, rappresentare, eseguire e recitare”. It is of course expected that these notions are equivalents, but it a matter of comparative law to assess whether they cover the same activities.

The Commons Deed carries a disclaimer, which is not very prominent, but still indicates that it doesn’t have any legal value: “The Commons Deed is not a license. It is simply a handy reference for understanding the Legal Code (the full license) — it is a human-readable expression of some of its key terms. Think of it as the user-friendly interface to the Legal

[55] Pallas-Loren, *op cit*, p. 19.

Code beneath. This Deed itself has no legal value, and its contents do not appear in the actual license."

Thus, only the Legal Code, a text of four to five pages, has legal value. The legal status of this three to four layers model will be further discussed more in detail (see section 4.1.2). As it will be later emphasized, the core clauses are less visible than the options, which are prominently advertised in the most accessible and visible formats of the licenses such as the title and the button. The Legal Code is deeply embedded under the notice button, two clicks away from the surface. First, the summary of the main freedoms and restrictions is accessible when clicking on the notice button. The notice sentence can be missing. The link embedded in the button/notice HTML is visible only when the mouse of the user goes on the button, otherwise the button appears as a static image. Some users may never click on, or even see, the summary of the provisions. And even once the user clicked on the link embedded in the notice button or sentence, the link to the actual text of the license is at the bottom of the summary, requiring again the user to scroll down until the last line of the Commons Deed: "This is a human-readable summary of the Legal Code (the full license) containing a link to the Legal Deed". The user seeing a notice button must thus spend quite some energy to access to the layer of the license which is said to have an actual legal value and this issue will be studied again to analyze its possible impact on the contract formation (section 4.1).

The Legal Code, which provisions will be described in details in the two next sections (the options in 2.2.2 and the core clauses 2.2.3), is a long text of 4 or 5 pages.

The machine-readable code is metadata which describes the license in the form of a digital rights expression. When selecting a license on the interface, it is possible to include additional information which "will be embedded in the HTML generated for (the chosen) license. This allows users of (the) work to determine how to attribute it or where to go for more information about the work". The fields are the following:

- The format of the work (audio, video, text, image, interactive, other),
- The title of the work,
- The name of the author or entity the licensor wishes the licensee to attribute,
- "The URL users of the work should link to. For example, the work's page on the author's site.",
- The URL of the source work if the work is derived from another work,
- A URL for more permission, where a user can obtain information about clearing rights that are not pre-cleared by your CC license.

This additional information can be embedded in the HTML code generated for the license, and will help to locate, identify and later manage the work. The machine-readable format allows search engines to index the work so that users may find works they can reuse. This is especially useful to support the remix culture and help locating works which can be copied or incorporated in larger works. Further applications could be developed to avoid inadequate or missing attribution and notice, and properly tag automatically derivative works with the appropriate licensing and attribution information. It is useful to be able to indicate not only the author or entity to be credited for attribution purposes. It would be even better to be able to also identify the licensor or rights owner, if different from the author or entity to be attributed.

The license code is attached to the work, and as we will see in the following section 2.2.3, the license requires the licensee to keep a link to, or a copy of the license when making copies or otherwise distributing or modifying the work. The persistence of the license code is therefore both needed and required by the license text. The machine-readable code, as a rights management information, is protected by anti-circumvention national legislations implementing WIPO Internet treaties. Such laws protect not only technical protection measures or DRMs against unauthorized circumvention, but also technical information measures against unauthorized removal[56]. On the top of the requirement expressed in the license to keep the licensing information with the work, it is an additional protection for the licenses which should stay attached to the works when they are further copied according to the freedoms expressed in the license. When a licensor attaches a CC license and additional copyright-related information to a work, the public is expected to keep that information intact when they share, modify and further distribute that work.

The importance of properly identifying the rights owner and of ensuring the license information will stay attached to the work will be analyzed in the section describing the legal validity of the agreement (section 4.1). After this description of the various layers or formats which constitute a CC license, we will present the other visible differences among licenses: the optional elements. They are displayed as icons in the button and as acronyms in the title.

[56] WIPO Copyright Treaty article 12 defines "rights management information" as " information which identifies the work, the author of the work, the owner of any right in the work, or information about the terms and conditions of use of the work, and any numbers or codes that represent such information, when any of these items of information is attached to a copy of a work or appears in connection with the communication of a work to the public ". This definition covers the machine-readable format of the CC licenses.

2.2.2 The license elements

"Full fat, semi-skimmed or no milk today"[57]? Creative Commons offers a flexible range of options for authors to distribute their works, between almost no control at all and a moderate or mild approach authorizing the public to copy the work without modifying it or making any profit. As we saw previously (section 2.1.2), the author is at the center of the system and gets to choose between many options, which offers a flexibility of choice. This large offer succeeds into gathering a large scope of authors with different needs and positions regarding the exercise of their exclusive rights and thus more works. But it also makes it difficult to assess what constitute the core freedoms of a CC license. It increases the information costs for both licensors and licensees in order to understand the differences between available options, and realize their long-term consequences.

In this section, we will present the license elements and their combinations, as well as the details and possible effects of the license elements provisions (BY, SA, NC and ND). License elements, or options, are the most visible component of the licenses. As we saw in section 2.2.1, they are the only elements of the licenses conditions which are accessible to the user in the visible formats of the system. The chosen combination constitute the title of the license, appearing in the notice button and at the top of the Commons Deed, the initials of the options are also in the Button and the icons representing the options are finally illustrating the text of the Commons Deed. Finally, they modulate the core grant expressed in main clauses which are less visible, and will be presented in the next section 2.2.3.

As explained in the introduction, the reference set of this study is constituted by the six licenses of the core suite in the current (3.0) unported version, in the legal code format. We will start by presenting the optional elements of these core licenses and then briefly present the other options or instruments which have been or still are available on the CC website: the Sampling suite, the Developing Nations license, the Founders Copyright, the Public Domain Dedication, CC0 and CC+.

After having presented the license elements (in this section 2.2.2), followed by the main clauses of the reference set (right after in section 2.2.3) and the legal functioning of these open content public licenses (section 2.3), we will be able to further identify in chapter 3 the sources of incompatibility within this reference set, with the other formats, versions, jurisdictions licenses of the CC system (e.g. from different, etc) and with other licenses of the open content ecosystem.

The six main licenses are combining four different elements, which authors can select online by answering the two following questions on a web interface, which is called the license chooser:

---

[57] Jones Richard, Cameron Euan, "Full fat, semi-skimmed or no milk today - creative commons licences and English folk music", *International Review of Law, Computers & Technology*, Volume 19, Issue 3 November 2005, p. 259-275. The authors use the milk metaphor to recall that Lawrence Lessig in *The Future of Ideas: The Fate of the Commons in a Connected World* "argues that intellectual property regimes need not be 'full on' (full fat) or 'full off' but partial (semi-skimmed). These ideas have found form in a more flexible regime of copyright through a series of alternative licensing contracts usually referred to as the Creative Commons licences."

License Your Work
With a Creative Commons license, **you keep your copyright** but allow people to copy and distribute your work provided they give you credit — and only on the conditions you specify here.

Allow commercial uses of your work?
Yes
No

Allow modifications of your work?
Yes
Yes, as long as others share alike
No

License your work: Creative Commons license chooser interface
Available at http://creativecommons.org/license/?lang=en

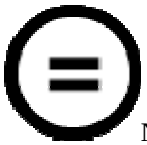
As described at http://creativecommons.org/about/licenses/, the CC licenses are a combination of one, two or three of the following four elements:


Attribution (BY)
Author lets others use her work if they give credit the way she requests.


Share Alike (SA)
The right holder allows others to make derivatives from your original work, but they should distribute these derivative works only under a license which is similar or recognized compatible to the license that governs your initial work.



Non-Commercial (NC)
The right holder let others use her work but for noncommercial purposes only. It does not mean that works can never be used for commercial purposes, but a separate license should be negotiated for commercial rights.


Non Derivative (ND)
The right holder authorizes others to copy, distribute, display, and perform only verbatim copies of her work, but does not grant the permission to make derivative works based upon it. The right to make adaptations can be licensed under a separate agreement.

The CC four license elements

The combination of the abovementioned license elements produces the six following licenses:

Attribution (BY)

This license lets others copy, distribute, display, perform and adapt the work, even commercially, as long as they credit the author of the original creation. This is the most permissive and accommodating of licenses offered, in terms of the broad scope of rights offered to others and minimal restrictions.

Attribution Share Alike (BY SA)

This license lets others copy, distribute, display, perform and adapt the work, even for commercial purposes, as long as they credit the author and license derivative creations of your work under identical terms. All new works will carry the same license, so any derivatives will also allow derivatives and commercial use. This license is often compared to open source software licenses, it maintains adaptations available under the same conditions.

Attribution Non-Commercial (BY NC)

This license lets others copy, distribute, display, perform and adapt the work for non-commercial purposes. Although their derivative works must also credit the author and be non-commercial, they don't have to license their derivative works on the same terms, meaning that derivatives can also be all rights reserved, unlike to those of the BY NC SA.

Attribution Non-Commercial Share Alike (BY NC SA)

This license lets others copy, distribute, display, perform and adapt the work in a non-commercial way, as long as they credit the author and license their derivatives under identical terms.

Attribution No Derivatives (BY ND)

This license permits redistribution in both commercial and non-commercial ways, as long the author is credited and the work copied or performed unmodified and in its integrality.

Attribution Non-Commercial No Derivatives (BY NC ND)

This license is the most restrictive of the six main licenses, allowing sole verbatim redistribution. This license is often called the "free advertising" license because it allows others to download works and share them with others as long as they attribute and link back to the author, but they can't reuse them in any way that would change them or use them commercially. The combination Attribution Non Commercial No Derivative Works only offers the possibility to copy and perform the work in limited circumstances[58].

The CC six core licenses

[58] It has once also carried the name of Music Sharing License and had a distinctive logo: http://creativecommons.org/choose/music

Let us now have a closer look at the legal code of the four license elements.

i. Attribution (BY)

The most liberal license, Attribution only. The Creative Commons Attribution license is used by the Open Access and the Open Educational Resources communities[59], which will gain more if works are reusable without restriction.

Attribution was an optional element in the initial version 1.0 of the licenses, one of the four optional elements which are presented in this subsection. It became a non-optional element and is featured in all the licenses, but it is still handled as an option or a License Element as far as the format is concerned because it appears in the title of the licenses, in the initials on the button and in the Commons Deed at the same level as the optional elements. Besides, the legal code of current version 3.0 also considers it in the same way it considers the optional elements[60]. Therefore, it is handled in this section at the same level as the other elements SA, NC and ND, even if it is not an option anymore.

This element answers a general concern of all creative communities: they agree to share their work, but only if they receive proper acknowledgement. What is understood as the legal norm in countries with moral rights appears to be a social norm in countries where authors are not always attributed. Beyond fame and pride, it is a common feeling among creators to share their creation only in exchange of public recognition, and perhaps more visibility on their other activities. But the clause sets up a standard of attribution which is higher than the legal and the social norms we are aware of. It is doubtful that is exercised to its fullest extent by licensors and implemented to its fullest extent by licensees.

The legal code related to the attribution element is long, detailed and difficult to access[61]. The

---

[59] The CC BY license complies with the definition of Open Access by the Budapest Open Access Initiative: “By "open access" to this literature, we mean its free availability on the public internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. The only constraint on reproduction and distribution, and the only role for copyright in this domain, should be to give authors control over the integrity of their work and the right to be properly acknowledged and cited.” http://www.soros.org/openaccess/read.shtml.

[60] In the Definitions section of the two Share Alike licenses, “"License Elements" means the following high-level license attributes as selected by Licensor and indicated in the title of this License: Attribution, ShareAlike/Attribution, Noncommercial, ShareAlike”.

[61] It is located in three subclauses, one subclause in the clause related to the license grant and two subclauses of the clause related to restrictions:
- In the license grant clause for the licenses authorizing adaptations to condition the exercise of this right to the identification of the changes made to the original work[61],
- In the second subclause of the restrictions clause as a positive obligation of the licensee to attribute the author or licensor as she requests, including the attribution of adaptations if they are authorized, and the way to exercise this obligation,
- And at the end of the first subclause of the restrictions (4.a.) as a negative obligation to remove upon request of the licensor such attribution from collections and adaptations to the extend they are authorized.

text varies between ND and non-ND licenses. The text is as follows, with the provisions related to derivatives italicized[62].

"If You Distribute, or Publicly Perform the Work or any Adaptations or Collections, You must, unless a request has been made pursuant to Section 4(a), keep intact all copyright notices for the Work and provide, reasonable to the medium or means You are utilizing:
(i) the name of the Original Author (or pseudonym, if applicable) if supplied, and/or if the Original Author and/or Licensor designate another party or parties (e.g., a sponsor institute, publishing entity, journal) for attribution ("Attribution Parties") in Licensor's copyright notice, terms of service or by other reasonable means, the name of such party or parties;
(ii) the title of the Work if supplied;
(iii) to the extent reasonably practicable, the URI, if any, that Licensor specifies to be associated with the Work, unless such URI does not refer to the copyright notice or licensing information for the Work; and
(iv) consistent with Section 3(b), in the case of an Adaptation, a credit identifying the use of the Work in the Adaptation (e.g., "French translation of the Work by Original Author," or "Screenplay based on original Work by Original Author").

(in clause 3 License grant)
to create and Reproduce Adaptations provided that any such Adaptation, including any translation in any medium, takes reasonable steps to clearly label, demarcate or otherwise identify that changes were made to the original Work. For example, a translation could be marked "The original work was translated from English to Spanish," or a modification could indicate "The original work has been modified.";

The credit required by this Section 4(c) *or 4(d)* may be implemented in any reasonable manner; provided, however, that in the case of a *Adaptation or* Collection, at a minimum such credit will appear, if a credit for all contributing authors of the *Adaptation or* Collection appears, then as part of these credits and in a manner at least as prominent as the credits for the other contributing authors.

For the avoidance of doubt, You may only use the credit required by this Section for the purpose of attribution in the manner set out above and, by exercising Your rights under this License, You may not implicitly or explicitly assert or imply any connection with, sponsorship or endorsement by the Original Author, Licensor and/or Attribution Parties, as appropriate, of You or Your use of the Work, without the separate, express prior written permission of the Original Author, Licensor and/or Attribution Parties.

If You create a Collection, upon notice from any Licensor You must, to the extent practicable, remove from the Collection any credit as required by Section 4(b), as requested.
If You create an Adaptation, upon notice from any Licensor You must, to the extent practicable, remove from the Adaptation any credit as required by Section 4(b), as requested."

The unported 3.0 legal code of the Attribution License Element

To sum up, the license foresees three provisions, "requested attribution", "unwanted attribution", and "non endorsement".

"Requested attribution" allows the licensor to require from the licensee a particular way to attribute the work by citing:
- The name of the author, licensor or any party,
- The title of the work,
- The source URL of the work,[63]
- For derivatives, a credit identifying the original author, the use of the original work and changes which have been made.[64]

The licensor may require these elements to be cited to the extent she supplies them, except for the last one because it is not possible. It is not quite clear how the licensee should fulfill this obligation in case no or insufficient information has been provided by the licensor who does

[62] We modified the layout of the clause in order to visually better separate the sentences, the language by itself being already difficult enough to read. We also modified the order of the 3 excerpts. It seems easier to present the subclauses in the logical order they are to be exercised, rather than in the order they are presented in the license, and thus to start with the requested attribution, including for adaptations, before the non endorsement and unwanted attribution requirements.
[63] But not the source URL of the original work for derivatives, which could be useful, as allowed by the Dublin Core field on the choser interface, see recommendations *infra* in chapter 5.
[64] This requirement may be difficult to express by the licensor and to achieve by the licensee. See recommendations of best practices *infra* in chapter 5 to lower the attribution requirements, by turning them into non-mandated best practices supported by automated applications performing the actual work of attribution properly.

not have or does not bother to put into practice the media literacy skills which are necessary to express this information. The standard of attribution is "a reasonable manner" except for Adaptations and Collections, where it should follow as a minimum the attribution standard of the other components[65].

"Non endorsement"
The licensee should not use the credit to imply that the author, licensor or party is endorsing the licensee or her use of the work. She "may only use the credit required by this Section for the purpose of attribution *in the manner set out above*", which is quite demanding.

"Unwanted attribution"
The licensee must be ready to remove the credit from Adaptations and Collections upon request from the licensor. This requirement raises practical questions. The licensor may never notice the work, or notice it late and make it impossible for the licensee to remove credits on works which have already been circulated, shared and reused.
Because this requirement seems to be related to the reputation of the author who might not want her name to be associated, we would suggest to cluster it, and maybe also the latter non-endorsement clause, with the moral rights provisions which comes right after in the license and will be studied in the section 2.2.3.

ii. Share Alike (SA)


The Share Alike option was inspired by the copyleft provision of the free and open source software licenses requiring derivatives to be licensed under the same terms. It will be compared with other open content licenses such as the GFDL and the FAL in section 3.5 of this study. It satisfies the needs of those who think that freedom must be preserved by requiring modifications to be shared with the same degree of freedom, in order to avoid a re-proprietarization of the commons. It is widely used for text-based creations and large ecosystems which need to be preserved from commercial appropriation. Attribution Share Alike can only be mixed with Attribution Share Alike, and Attribution Non-Commercial Share Alike only with Attribution Non-Commercial Share Alike.

The Share Alike text presented below appears in the restriction clause of the license[66]:

[65] The compliance to this requirement may be difficult to assess.
[66] For formatting reasons, we reorganized the text of the clause, removed a substantial portion at the center of the clause and added the definitions of CC Compatible License and License Elements which appear in the Definition section.

"You may Distribute or Publicly Perform an Adaptation only under the terms of:
(i) this License;
(ii) a later version of this License with the same License Elements as this License;
(iii) a Creative Commons jurisdiction license (either this or a later license version) that contains the same License Elements as this License (e.g., Attribution-NonCommercial-ShareAlike 3.0 US);
(iv) a Creative Commons Compatible License.
(…)
"Creative Commons Compatible License" means a license that is listed at http://creativecommons.org/compatiblelicenses that has been approved by Creative Commons as being essentially equivalent to this License, including, at a minimum, because that license:
(i) contains terms that have the same purpose, meaning and effect as the License Elements of this License; and,
(ii) explicitly permits the relicensing of adaptations of works made available under that license under this License or a Creative Commons jurisdiction license with the same License Elements as this License.
(…)
"License Elements" means the following high-level license attributes as selected by Licensor and indicated in the title of this License: Attribution, ShareAlike.
(…)
This Section 4(b) applies to the Adaptation as incorporated in a Collection, but this does not require the Collection apart from the Adaptation itself to be made subject to the terms of the Applicable License."

The unported 3.0 legal code of the Share Alike License Element

The Share Alike language is relatively clear. It states that adaptations must be licensed under the same terms as the original work, and it defines what terms are declared compatible: the same BY SA license, a later version of the BY SA license, a jurisdiction version of the same or a later version of the BY SA license.

We will further discuss the possible impact of the clause, which declares compatible licenses the texts of which are different:

- Subsequent versions may contain different terms (section 3.2),
- Jurisdictions versions contain different terms (section 3.4) and
- Other open content licenses have different terms (section 3.5)

and seems to bind licensors and licensees to obligations they are not aware of and could not consent to (section 4.1).

iii. Non Commercial (NC)

The Non Commercial option restricts the exercise of the rights granted by the license to non-commercial situations. In other words, the licensor reserves commercial rights.

"You may not exercise any of the rights granted to You in Section 3 above in any manner that is primarily intended for or directed toward commercial advantage or private monetary compensation. The exchange of the Work for other copyrighted works by means of digital file-sharing or otherwise shall not be considered to be intended for or directed toward commercial advantage or private monetary compensation, provided there is no payment of any monetary compensation in connection with the exchange of copyrighted works."

The unported 3.0 legal code of the Non Commercial License Element

This provision has been widely criticized. It is not an acceptable restriction for the copyleft and Free/Libre and Open Source Software communities: it prevents the definition of a clear freedom for the community and it may even be counter-productive[67]. "The people who are likely to be hurt by an -NC license are not large corporations, but small publications like weblogs, advertising-funded radio stations, or local newspapers." The Share Alike clause could be a better alternative: "while not applicable to monetary benefits, (it) does protect the content from abusive exploitation without forbidding experiments (…) Any company trying to exploit your work will have to make their "added value" available for free to everyone. The company does not, however, need to share the income from the "added value". Seen like this, the "risk" of exploitation turns into a potentially powerful benefit depending on the value added to the content." [68]

Besides, even if the clause text is less legalese than other provisions, it leads to a lot of confusion[69] and doubts related to its interpretation causes legal uncertainty. The first common misunderstanding, coming from people who may not have read the clause but only interpreted the notice button or title format, is that it would prevent licensors from making any profit. It is not the case, the restriction applies to uses made by licensee, not the licensor. In the same vein, some think that licensors (or licensees) have to be non-profit institutions, which is also not true. Once it has been clarified that the provision targets uses by the licensee, the scope of the definition "*primarily intended for or directed toward commercial advantage or private monetary compensation*" remains open to legal interpretation. The line between commercial and non-commercial uses is thin and leads to categorization difficulties.[70] Unlike the concept of attribution and derivative work, the notion of non-commercial use is not defined by copyright legislations. In the United States, it is cited by the law as a factor to determine whether a situation can be considered as fair use[71]. A strict interpretation reduces the possibility a work will be actually reused beyond straightforward cases, such as a personal website without advertising banners, or a class in a public school. However, the element was chosen by three quarters of the licensors in 2004 and more than half of the licensors in 2006,[72] expressing a concern that others may make profit of one's work while one was unable or

---

[67] See Chen and Mako-Hill, *op cit*, see also reasons for not using NC by Möller Eric, "Creative Commons -NC Licenses Considered Harmful", *Kuro5hin*, September 2005, which evolved into the editable paper "The Case for Free Use: Reasons Not to Use a Creative Commons-NC License", 2005-2007.
http://www.kuro5hin.org/story/2005/9/11/16331/0655 and http://freedomdefined.org/Licenses/NC.

[68] Möller, *ibid*.

[69] The reading of the Australian Copyright Council may be incorrect according an author citing Australian Copyright Council, Information Sheet G094: Creative Commons Licenses (May 2006): Weatherall Kimberlee, "Would you ever recommend a Creative Commons license?", Unlocking IP 2006 Conference, "Creating Commons: The Tasks ahead in Unlocking IP," UNSW AGSM, Sydney, 10-11 July 2006, Australasian Intellectual Property Law Resources (AIPLRes) 22, 2006.
http://www.austlii.edu.au/au/other/AIPLRes/2006/22.html

[70] "For example, a recurrent question in the educational context, and one of the most debated, is whether the NC restriction allows a user to charge for copying and distributing the licensed material and for associated overhead expenses including salaries, irrespective of the user's business status (non-profit, for-profit, government). Some believe that the for-profit status of the business itself should preclude this; others disagree." In Rutledge Virginia, "Fair Comment: Towards a Better Understanding of NC Licenses", Commonwealth of learning, *Connections*, February 2008. http://www.col.org/news/Connections/2008feb/Pages/fairComment.aspx

[71] "1. the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes", Copyright Act of 1976, 17 U.S.C. § 107

[72] On choice of options, see Cheliotis Giorgos, "Creative Commons Statistics from the CC-Monitor Project", *presentation at the iCommons Summit*, Dubrovnik, June 14-17, 2007.
http://creativecommons.org/weblog/entry/7551

unwilling to do so. Therefore, even if this option is limiting the reuse of works because it is difficult to assess whether a usage is truly non-commercial, it largely contributed to the success of the movement in terms of popularity within the general public.

A study on "Defining Non Commercial" has been carried out by CC and a report has been published in 2009 based on a market research among users[73]. One of the most interesting findings is that in many cases, licensees have a stricter interpretation of what uses constitute a commercial use than licensors, whose expectations should therefore be met. It will be later evaluated:

> "If the better approach might be to adopt a "best practices"[74] approach of articulating the commercial/noncommercial distinction for certain creator or user communities apart from the licenses themselves. (…) While the costs of license proliferation are already widely appreciated and resisted by many, the study weighs against any lingering temptation to offer multiple flavors of NC licenses due to strong agreement on the commerciality of certain use cases that, in the past, may have been considered by some to be good candidates for splitting off into specialized versions of the NC term, such as online advertising."[75]

Despite the legitimate critiques of the NC option, it should be recognized that it intends to support many business models (online advertising such as banners on a website, selling of a physical support such as a compilation or a book, illustration of a commercial, etc) and its potentiality should not be neglected, especially for the music and book industry. It also clarifies the situation of file-sharing and private remixing by explicitly authorizing these practices[76], while reserving possible remuneration on commercial uses, such as the collection of royalties from public performance. We will see later[77] that this model has the potential to be accommodated by collective management societies who may collect royalties on commercial uses. However, the model is not sustained and even jeopardized by incompatibilities with the current collective management statutes and practices of many collecting societies.

---

[73] Creative Commons, " Defining "Noncommercial". A Study of How the Online Population Understands "Noncommercial Use" ", September 2009. The study report, data and excerpts from the executive summary can be accessed at http://wiki.creativecommons.org/Defining_Noncommercial. See also the blog announcement available at http://creativecommons.org/weblog/entry/17127

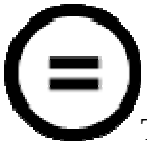
[74] We will discuss this approach in the final section of this study. Guidelines have already been published on the CC website http://www.creativecommons.se/NonCommercialGuidelines.pdf
and by MIT : http://ocw.mit.edu/OcwWeb/web/terms/terms/index.htm#noncomm
On the issues raised by guidelines, which interpretation might differ from interpretation by Courts, see also the Criticism of the non-commercial clause by the OER Africa.
http://www.oerafrica.org/CriticismsoftheNonCommercialClause/tabid/873/Default.aspx

[75] Creative Commons, " Defining "Noncommercial". A Study of How the Online Population Understands "Noncommercial Use" ", *op cit*, p. 77.

[76] File-sharing is a practise which has been criminalized in many countries while its negative impact on sales is not demonstrated. Thus, the NC clause brings legal certainty and security to musicians audience. " The decision by CC to exclude this specific use case in its noncommercial licenses was driven in part by the *Napster* court decision, in which the court concluded that the trading of music online was commercial in nature even though no money exchanged hands. A&M Records, Inc. v. Napster, Inc., 239 F.3d 1004 (9th Cir. 2001). ", in Creative Commons, " Defining "Noncommercial ", *ibidem*, p. 17.

[77] See *supra* sections 3.4 for the differences among jurisdictions and 4.2 for an analysis of the impact of the incompatibility of some collecting societies statutes with all the CC licenses.

iv. No Derivative (ND)



The No Derivative license element caters to the needs of those who do not want their work to be modified. However, it will not prevent its aggregation in a collection, changes of formats, nor modifications which are authorized by jurisdictions' exceptions and limitations, such as parody, or transformative use, a fair use factor. It answers to fears of being associated with works one would not approve of or having one's ideas mutilated or distorted. Some authors choose this option without realizing that it will prevent some use cases they would support, such as the translation of their scientific article in a foreign language, or the illustration of a documentary with their music. Perhaps they have reputation concerns and do not realize that also the non-ND licenses contain a clause asserting moral rights, require authors of derivatives to describe their adaptation and prevent licensees to claim any association or endorsement by the author of the original work as we just saw in the Attribution clause description. A line must be drawn between integrity and the right to make derivatives. The ND clause should not be used for the sole purpose of ensuring the integrity of the work and the non-endorsement of the adaptation.

The ND option does not actually correspond to a clause *per se* in the license. By contrast, the non-ND licenses have additional clauses, in the form of a broader license grant in clause 3. ND licenses authorize:

"a. to Reproduce the Work, to incorporate the Work into one or more Collections, and to Reproduce the Work as incorporated in the Collections; and,
b. to Distribute and Publicly Perform the Work including as incorporated in Collections."

While non-ND licenses authorize the making of adaptations, and the difference is italicized below:

"a. to Reproduce the Work, to incorporate the Work into one or more Collections, and to Reproduce the Work as incorporated in the Collections;
b. to create and Reproduce Adaptations provided that any such Adaptation, including any translation in any medium, takes reasonable steps to clearly label, demarcate or otherwise identify that changes were made to the original Work. For example, a translation could be marked "The original work was translated from English to Spanish," or a modification could indicate "The original work has been modified.";
c. to Distribute and Publicly Perform the Work including as incorporated in Collections;
d. and, to Distribute and Publicly Perform Adaptations."

There are finally two other differences between ND and non-ND licenses which will both be later analyzed[78]:

- In the format clause, to explain that the right to make modifications which are technically necessary does not include the right to make adaptations, and
- In the moral rights clause, to confirm that adaptations must not be prejudicial to the author's honor or reputation.

"The above rights may be exercised in all media and formats whether now known or hereafter devised. The above rights include the right to make such modifications as are technically necessary to exercise the rights in other media and formats, *but otherwise you have no rights to make Adaptations*." (end of clause 3.)

"Except as otherwise agreed in writing by the Licensor or as may be otherwise permitted by applicable law, if You Reproduce, Distribute or Publicly Perform the Work either by itself or as part of any *Adaptations or* Collections, You must not distort, mutilate, modify or take other

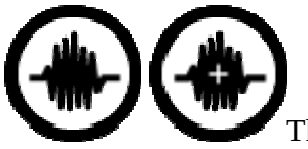
[78] See *infra* section 2.2.3.

derogatory action in relation to the Work which would be prejudicial to the Original Author's honor or reputation. *Licensor agrees that in those jurisdictions (e.g. Japan), in which any exercise of the right granted in Section 3(b) of this License (the right to make Adaptations) would be deemed to be a distortion, mutilation, modification or other derogatory action prejudicial to the Original Author's honor and reputation, the Licensor will waive or not assert, as appropriate, this Section, to the fullest extent permitted by the applicable national law, to enable You to reasonably exercise Your right under Section 3(b) of this License (right to make Adaptations) but not otherwise.*" (last sub-clause of clause 4.)

Reserving modifications does not encourage creativity and reappropriation. Moreover, it prohibits translation. Exercising some control on adaptations can be already be achieved through the BY, the SA and the NC license elements. The BY clause requires the licensee author of an adaptation to identify the modifications from the original work and contains the non-endorsement provision to protect the original author. The SA clause constraints the terms under which adaptations may be released. The BY NC and the BY NC SA licenses authorize modifications, but not if the derivatives are used in a commercial way. As we discussed, the BY NC SA combination, the most popular of all the CC licenses, may satisfy those supporting the sharing and the remix culture, but who are reluctant to see others succeeding at making profit of one's work.

v. Instruments outside the core suite

Besides the BY, SA, NC and ND license elements constituting the licensing core suite, other licenses or tools have been made or are still available on the CC website: the Sampling suite, the Developing Nations license, the Founders Copyright, the Public Domain Dedication, CC0 and CC+. Here is a brief description of these instruments.



The Sampling licenses "let artists and authors invite other people to use a part of their work and make it new."[79] The interface[80] to select these licenses is not easily accessible from the CC website anymore. It was anyway not widely used even when the choice was offered next to the standard interface. Only three jurisdictions (Brazil, Germany and Taiwan) ported these licenses, which will not be further studied.

The Sampling 1.0 license[81] was retired because it did not allow reproducing the entire work even for non-commercial purposes[82]. It would only allow to use the work partially or non substantially or transform it substantially through employing "sampling," "collage," "mash-up," or other comparable artistic technique".

The three Sampling licenses all prohibit the reuse for "advertising and promotional uses", "except for advertisement and promotion" of the new work.

The Sampling + 1.0 license[83], in addition to allowing the making of the partial kind of derivative works which has been described just above in the Sampling license, also allows "noncommercial sharing of verbatim copies", thus "+" as the core grant common to all the CC licenses (at minimum BY NC ND) is being added to the Sampling right.

---

[79] http://creativecommons.org/about/sampling
[80] http://creativecommons.org/choose/sampling
[81] http://creativecommons.org/licenses/sampling/1.0/legalcode and http://creativecommons.org/licenses/sampling/1.0/
[82] http://creativecommons.org/retiredlicenses : "it did not permit non-commercial verbatim sharing"
[83] http://creativecommons.org/licenses/sampling+/1.0/legalcode

The NC Sampling + 1.0 license[84] grants the same rights than the Sampling + license, except that not only the verbatim copies are submitted to the NC provision, but also the derivative work resulting from the sampling activity, which is called "Re-creativity right" in all these licenses and correspond to a portion only of the right to make Derivative works granted in the non-ND licenses of the core suite.
The rights of the Sampling licenses vary substantially from the usual CC legal texts and are therefore difficult to understand.

The Developing Nations 2.0 license authorizes commercial use and the making of derivatives in developing nations, and therefore does not contain the text of the clauses SA, NC or ND. The exercise of the rights are however submitted to a specific provision displayed at the end of the Restriction clause 4, stating that only Developing Nations can access the work:

"c. The Work and any Derivative Works and Collective Works may only be exported to other Developing Nations, but may not be exported to countries classified as "high income" by the World Bank.
d. This License does not authorize making the Work, any Derivative Works or any Collective Works publicly available on the Internet unless reasonable measures are undertaken to verify that the recipient is located in a Developing Nation, such as by requiring recipients to provide name and postal mailing address, or by limiting the distribution of the Work to Internet IP addresses within a Developing Nation."

The Developing Nations 2.0 license was also retired because only a restricted audience was authorized to copy the work, while other users located in developed countries were not allowed to reproduce the work, even for NC purposes[85].

After the Sampling and the Developing Nations licenses, which are not fulfilling requirements of legal standardization and harmonization of a core grant, another series of tools deserves a short presentation: the public domain tools (Founders Copyright, Public Domain Certification, CC0) and the CC+ protocol. They differ from the standard suite not only substantially, but also procedurally: compared to the standard user interface, they all require explicit consent from the prospective licensor who is prompted to provide more information such as the name of the author.

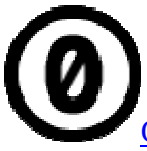
The Founders Copyright[86] allows putting a work in the Public Domain 14 years after its creation, reducing thus the exercise of copyright to the duration which had originally been foreseen in 1790. It may be seen as a small-scale experiment of re-establishing formalities. "To re-create the functionality of a 14- or 28-year copyright, the contributor will sell the copyright to Creative Commons for $1.00, at which point Creative Commons will give the contributor an exclusive license to the work for 14 (or 28) years."[87] Unlike to the other licenses of the CC system, it targets US authors only, they transfer their rights to CC who provides an online registry and requires filling a form[88] to which CC will provide an answer. In particular, the applicant is asked to provide the name of the copyright holder and, in order to secure she represents the rights which will be exercised by CC, to answer by yes or no to the following questions:
"Do you have exclusive rights to this work?
Are there parts of your work that are from other sources (quotes, pictures, etc.)?
Is this a derivative work? (includes translations)"

[84] http://creativecommons.org/licenses/nc-sampling+/1.0/legalcode
[85] http://creativecommons.org/retiredlicenses
[86] http://creativecommons.org/projects/founderscopyright/
[87] http://creativecommons.org/projects/founderscopyright/
[88] http://creativecommons.org/projects/founderscopyright/inquiry

The Copyright-Only Dedication or Public Domain Certification[89] is used to certify a work that is already in the public domain. Unlike the standard licenses, obtaining the legal code[90] requires the user to explicitly manifest and express her consent to a text, which corresponds to the text of the license[91] by clicking a box[92]:

"I have read and understand the terms and intended legal effect of this tool, and hereby voluntarily elect to apply it to this work."

In addition to the main licenses, two additional tools have been recently developed: CC+ and CC0.



CC0 (CC "Zero") is a waiver of copyright, neighboring and related rights, and sui generis rights. CC0 is intended to facilitate access to and reuse of works by placing them as nearly as possible into the public domain before applicable copyright term expires. CC0 can be used for all kinds of works and also for non-copyrightable scientific data sets, or databases of works in the public domain curated by libraries, museums or archives. CC0 is a "no rights reserved" option. CC recommends[93] using CC0 instead of the Public Domain Certification for works which are still protected by copyright. Even if there is no registration process, the user is also prompted[94] to provide name, URL, title, territory and to manifest her consent:

"I hereby waive all copyright and related or neighboring rights together with all associated claims and causes of action with respect to this work to the extent possible under the law."
"I have read and understand the terms and intended legal effect of CC0, and hereby voluntarily elect to apply it to this work."

A double-click confirmation is even required:

---

[89] http://creativecommons.org/licenses/publicdomain/
[90] http://creativecommons.org/licenses/publicdomain/
[91] Confirm Your Public Domain Certification
Copyright-Only Dedication (based on United States law) or Public Domain Certification
The person or persons who have associated work with this document (the "Dedicator" or "Certifier") hereby either (a) certifies that, to the best of his knowledge, the work of authorship identified is in the public domain of the country from which the work is published, or (b) hereby dedicates whatever copyright the dedicators holds in the work of authorship identified below (the "Work") to the public domain. A certifier, moreover, dedicates any copyright interest he may have in the associated work, and for these purposes, is described as a "dedicator" below.
A certifier has taken reasonable steps to verify the copyright status of this work. Certifier recognizes that his good faith efforts may not shield him from liability if in fact the work certified is not in the public domain.
Dedicator makes this dedication for the benefit of the public at large and to the detriment of the Dedicator's heirs and successors. Dedicator intends this dedication to be an overt act of relinquishment in perpetuity of all present and future rights under copyright law, whether vested or contingent, in the Work. Dedicator understands that such relinquishment of all rights includes the relinquishment of all rights to enforce (by lawsuit or otherwise) those copyrights in the Work.
Dedicator recognizes that, once placed in the public domain, the Work may be freely reproduced, distributed, transmitted, used, modified, built upon, or otherwise exploited by anyone for any purpose, commercial or non-commercial, and in any way, including by methods that have not yet been invented or conceived.
[92] http://creativecommons.org/choose/publicdomain-2
[93] http://creativecommons.org/publicdomain
[94] http://creativecommons.org/choose/zero/waiver

"Are you certain you wish to waive all rights to your work? Once these rights are waived, you cannot reclaim them."

Then, a Commons Deed[95] and a Legal Code[96] are available as usual after having selected the License Elements of the standard interface.

CC+ (CC "Plus") is not an additional license, but a technology to signal the addition of more rights beyond a CC license grant, for instance to clear commercial rights, or to obtain more warranties, and indicate the link to these additional permissions. It has a strong potential but it is not advertised on the license chooser, thus not accessible for the average user of the system.

Finally, if license options are to be defined as license elements or features which have an icon, we should mention non-CC licenses which have a CC wrapper (machine-readable metadata and human-readable Commons Deed), namely the GNU-GPL and GFDL as well as the BSD[97], which conditions even have ad-hoc icons (notice, source code, no endorsement) which could be reused in the actual CC licenses human-readable format.

[95] http://creativecommons.org/publicdomain/zero/1.0/
[96] http://creativecommons.org/publicdomain/zero/1.0/legalcode
[97] http://creativecommons.org/licenses/GPL/2.0/
http://creativecommons.org/licenses/LGPL/2.1/
http://creativecommons.org/licenses/BSD/

2.2.3 The main clauses

We presented all the available options of the CC system in the previous section, with a focus on the license elements which are deployed around the licenses main clauses. We will now analyze the detail of these main clauses. In order to provide legal certainty and security, it matters to find out what is exactly covered and whether it is made clear to the user.

The core grant of the CC system is an authorization to copy, display, perform and distribute the work without modifying it and for non-commercial purposes only, to which more freedoms can be granted when playing with the license elements. The user interface in the CC Lab[98], a section of the CC website dedicated to experimental projects, makes it possible to play with the license elements in another way than the usual license chooser interface[99], making it cognitively easier to understand that the main clauses express positive rights from which the NC and ND options are taking away.

The license elements play a very important role in the CC system, they appear even before the rights they alter. It may seem illogical to present conditions pertaining to rights before rights themselves. However, the license elements are accessible before the main clauses in the license chooser interface, in the notice button and in the title of the license. The main clauses appear only in the deeper layer, the Legal Deed, and to a lesser extend in the Commons Deed a summarized version deprived of legal value.

The license elements, which are very visible in the Notice Button and the Commons Deed may be hiding the substance of the license to the user, who must read the main clauses behind the options. Besides information costs, the question is whether these main clauses are not only visible, but also clear substantially to the user. Knowing precisely which rights are granted by whom on which subject matter is essential for the validity and the coherence of the system.

We will describe systematically the main provisions of the eight clauses of a CC license in its unported 3.0 version. This presentation will allow us to clarify what is the subject matter, and to compare the core grant of the 3.0 unported license legal deed[100] with the other licenses versions, jurisdictions and formats, in order to identify differences and potential sources of incompatibilities in chapter 3. Most of the core grant is not mentioned in the Commons Deed and therefore not very accessible to the average user, who is nevertheless expected to consent to the legal code (section 4.1.2).

The six main Creative Commons licenses authorize as a minimum to copy, perform and distribute the unmodified work for free, provided that the original author is properly attributed and that no direct remuneration is perceived in exchange for the work. The licenses' optional elements NC, ND and SA specify the nature of this core grant and prescribe whether works

---

[98] The user can play with the bricks of a license on the Freedoms License Generator available in the ccLab at http://labs.creativecommons.org/demos/freedomslicense/. This license engine is presented as a puzzle and may have different cognitive results on the understanding by the user than the usual license choser interface: "Not all combinations are possible, but as you experiment with the selections, you can see the different licenses that result."

[99] http://creativecommons.org/choose

[100] http://creativecommons.org/licenses/by-nc-nd/3.0/legalcode

can be used for commercial purposes, may be adapted and if yes, how such adaptations may be redistributed. All the CC licenses authorize the public to copy and distribute the work, including in collective works, and to display or perform it in all media and formats, including digital file-sharing. As we noticed in the previous section describing the license elements, the six core licenses are an assembling/assembly of clauses which vary according to the combination.

Methodologically, in order to analyze all the main clauses, we have to examine the squeletton of a license, e.g. the core provisions without the license elements and without the small textual variations between ND and non-ND licenses, depending on whether adaptations are authorized (variations which have been identified in italic in the previous section 2.2.2.). We cannot simply analyze the core freedoms expressed in the most restrictive or the most liberal licenses (the BY NC ND license or the BY license), neither can we use the license used during the porting process because it contains all the clauses (the BY NC SA license).

We will compare systematically the text of the definitions and the main clauses with definitions provided in the latest versions of the international conventions which are cited in article 8f[101]: "The rights granted under, and the subject matter referenced, in this License were drafted utilizing the terminology of" the Berne Convention for the Protection of Literary and Artistic Works (hereafter the Berne Convention), the Rome Convention for the Protection of Performers (the Rome Convention), the WIPO Copyright Treaty (WCT), the WIPO Performances and Phonograms Treaty (WPPT). Before analyzing the compatibility between licenses, the compatibility with international law is to be checked in order to detect possible inconsistencies or confirm that the system is viable. Of course the licenses do not have to mention all the notions of the international conventions and can go beyond, but it is important to check what notions are exactly covered to make sure that no right or party has been left out. Indeed, the grant appears to be as broad as possible and it can therefore be expected that all works and all rights are addressed by the licenses and that they are no hidden restrictions on the nature of works and rights covered. International conventions have been chosen as a standard to assess the licenses not because they are a model or because they are inclusive, but because the unported version has been drafted according to them, and because as international law instruments, they are a minimum to be implemented in national legislations of their member states. It should be noted that not all countries are members of all conventions, indeed the United States are not a contracting party of the Rome Convention; thus, including its provisions in the license text goes beyond minimum standards.

The license consists of a foreword and eight clauses (as well as a header and a notice with information about CC as a corporation). These provisions may also be found in other open content licenses:

- Definitions of items covered (what is a work) and parties involved (licensor, author…),
- The exact nature of the rights granted,
- The restrictions that may apply to the grant, including the BY, SA, NC and ND restrictions previously described in section 2.2.2,

[101] The Universal Copyright Convention is also cited in article 8f, nevertheless no parallel between the definitions of the licenses and of this convention has been found.

- Some procedural requirements accompanying works copies and performances: a license notice must be conveyed with the work, which author and original work in case of derivatives must be credited in an appropriate way as we saw in section 2.2.2 developments related to attribution,
- The relationship with applicable law: the licenses apply in addition to the law, and in particular they claim to not conflict with exceptions to exclusive rights, moral rights and compulsory licensing schemes in the jurisdictions where they exist, therefore, they may yield in front of incompatible legal provisions which may be unknown from the licensor,
- Exclusion of representations and warranties and limitation of liability,
- Other standard clauses, such as:
    - The termination of the license for those licensees who do not comply with the terms of the license, leading to the return to an all-rights-reserved scenario and possibly copyright infringement if the use does not stop,
    - The possibility for the licensor to stop distributing the work under the license does not lead to withdrawing rights which have been granted to licensees prior to this decision, providing legal security to those who have already copied, distributed or otherwise incorporated the work in their own creation.

The text starts with a foreword, stating that the use of the work is governed by the license as well as applicable law. We will see in greater detail in section 2.3.1 how an agreement can be formed between the parties. Let us now analyze the main clauses one by one.

i. Definitions

The license starts with definitions of the subject-matter (Work, Adaptation, Collection), the rights (Reproduce, Distribute, Publicly Perform) and the parties involved (Licensor, Original Author, You). We will present them in the order of their usage, not in alphabetical order as it is the case in the license. We will use the capital letter further in this study when exactly referring to these notions as defined by the license.

a. Work

The CC definition for Work comes from the Berne Convention Article 2.1. “Literary and artistic works”, with few variations:

> **"Work"** means the literary and/or artistic work offered under the terms of this License including without limitation any production in the literary, scientific and artistic domain, whatever may be the mode or form of its expression including digital form, such as a book, pamphlet and other writing; a lecture, address, sermon or other work of the same nature; a dramatic or dramatico-musical work; a choreographic work or entertainment in dumb show; a musical composition with or without words; a cinematographic work to which are assimilated works expressed by a process analogous to cinematography; a work of drawing, painting, architecture, sculpture, engraving or lithography; a photographic work to which are assimilated works expressed by a process analogous to photography; a work of applied art; an illustration, map, plan, sketch or three-dimensional work relative to geography, topography, architecture or science; a performance; a broadcast; a phonogram; a compilation of data to the extent it is protected as a copyrightable work; or a work performed by a variety or circus performer to the extent it is not otherwise considered a literary or artistic work.

Berne definition refers to the expression “literary and artistic works” and uses the plural, while CC designates the literary and/or artistic work and provides the examples of the Berne

Convention in the singular and adding "without limitation" and "including digital form" in order to not exclude other forms not mentioned in the license definition[102]. CC also adds "a performance; a broadcast; a phonogram; a compilation of data to the extent it is protected as a copyrightable work; or a work performed by a variety or circus performer to the extent it is not otherwise considered a literary or artistic work". However, CC does not include the first fixation of a film or broadcast, while videograms are targeted by Berne article 9.3: "Any sound or visual recording shall be considered as a reproduction for the purposes of this Convention" and broadcasts are addressed by Rome article 3f: "sounds or of images and sounds".

Performance, broadcast and phonogram are not defined, but performers, broadcasters and producers of phonograms are defined in another definition, as in the Rome Convention. "Variety and circus artists (…) who do not perform literary or artistic works ", thus a slightly different phrasing, are mentioned under Rome Convention article 9.

As we will see in the definition of Original Author, and like in article 2.a. of the WPPT, the CC indirect definition of performer includes the performance of literary or (and not "and) artistic work, but also the performance of expressions of folklore, which are not copyrightable works by themselves.
"A compilation of data to the extent it is protected as a copyrightable work" most likely targets compilations as defined at article 5 of the WCT "Compilations of data or other material, in any form, which by reason of the selection or arrangement of their contents constitute intellectual creations", but does not formally encounter compilations of other material, for instance compilations of copyrightable works as opposed to compilations of non-copyrightable data.
The fixation of a musical composition, the phonogram, is mentioned, but the fixation of a film and the fixation of a broadcast is not mentioned, while "visual or audio-visual fixation" is indirectly mentioned in Rome Convention article 19[103].

Even if the use of the expression "without limitation" and "including digital form" limits the risk to leave out forms of expressions, there are several uncertainties in the subject matter, namely what is a compilation of data and whether compilations of works, databases and first fixations of films and broadcasts are covered, to the extend they are neither "compilations of data protected as copyrightable works" nor "cinematographic works" or "broadcasts" as targeted by the definition of Work. It would be preferable to include explicitly first fixations of films and broadcasts in order to be sure they are also covered. Indeed, we cannot assume that they have been intentionally left out of the scope of the licenses. We will further come back to the question of databases in section 4.2.2.; unlike to videograms, CC as an organization expressed at some point the intention to exclude databases of the licenses scope, as they are not a subject matter of copyright per se in many countries.

Also, in case the item targeted by the license is a complex work which combines several forms of expression, such as a musical composition, a performance and a phonogram, they

[102] Such as for instance "official texts of a legislative, administrative and legal nature", "a matter for legislation in the countries of the Union to determine the protection to be granted to", Berne Convention article 2.4.
[103] And covered by the acquis communautaire (Rental Directive, EUCD).

should all be covered[104]. It could be made clearer that the Work can include several types of Works, e.g. a work and its performance and its fixation, in the case of a music title, all the more as users are not defining specifically enough the Work in their License Notice.

In previous versions of the licenses, Work was defined as "the copyrightable work of authorship". Version 3.0 aims at grounding the text of the licenses in international law rather than in American law. However, the definition of what is a protected work under copyright or which items are protected by neighboring or sui generis rights is a matter for legislations in the country. It is also questionable whether related rights are part of the category of "copyright" (as it is the case for its equivalent of Literary and Artistic Property for instance in France) or if they should be mentioned explicitly and separately. The latter option probably provides more certainty. Therefore, Work could have been defined as "the copyrightable of work of authorship and/or the other forms of creation protected by related rights". Otherwise, in the case of a CD for instance, the underlying work, the musical composition, could be CC-licensed, but neither the performance nor the phonogram.

b. Adaptation

> **"Adaptation"** means a work based upon the Work, or upon the Work and other pre-existing works, such as a translation, adaptation, derivative work, arrangement of music or other alterations of a literary or artistic work, or phonogram or performance and includes cinematographic adaptations or any other form in which the Work may be recast, transformed, or adapted including in any form recognizably derived from the original, except that a work that constitutes a Collection will not be considered an Adaptation for the purpose of this License. For the avoidance of doubt, where the Work is a musical work, performance or phonogram, the synchronization of the Work in timed-relation with a moving image ("synching") will be considered an Adaptation for the purpose of this License.

The first part of the CC definition for Adaptation comes from the Berne Convention definition for Derivative works "Translations, adaptations, arrangements of music and other alterations of a literary or artistic work", except that derivative work is not the name of the category but inserted within the list. It includes the adaptations of works, performances and phonograms, but not the adaptation of broadcasts. Therefore, there would be the risk to not authorize the adaptation of a broadcast licensed under a non-ND license, while the Licensor who wouldn't have read the clause would probably intend to authorize it, and the Licensee would probably not be aware it is not included.

It also "includes cinematographic adaptations or any other form in which the Work may be recast, transformed, or adapted including in any form recognizably derived from the original" and the synchronization of the work when it is music on moving images. The two latter provisions are not based on international conventions but on US law, the first is an except from the US Copyright Act section 101 definition for Derivative work. Qualifying the synchronization of musical works on moving images as a Derivative work and not as a Collective work is a common practice. Synchronization on a movie, a TV programme or advertisement usually involves modifications such as cuts of the original work. Based on questions and discussions on the CC mailing lists as well as infringement cases, many users are not aware of the fact that synchronization is considered an adaptation. They do not realize that the Share Alike provision applicable to a music track should be transmitted to the moving

[104] See Angelopoulos Christina, "Creative Commons and Related Rights in Sound Recordings: Are the Two Systems Compatible?", Institute for Information Law, December 2009, 44 p. http://www.ivir.nl/publications/angelopoulos/Creative_Commons_and_Sound_Recordings.pdf

images that would embed the music. This creates legal insecurity if the provision is ignored.[105]

It is a possibility that in the future, Adaptations may be defined in a broader way, in order to include more modifications and make Share Alike stronger and applicable to more Works, for instance qualify the incorporation of an image into a text as an adaptation. This point has been discussed at the occasion of a statement of intent regarding compatibility with the GFDL license[106].

c. Collection

The CC definition for a Collection comes from the definition of a Collection in the Berne Convention article 2(5). It encompasses not only works, but also performances, phonograms or broadcasts, but as noticed above, it does not mention videograms to the extend they are a different instantiation of a cinematographic work or a broadcast and this could be corrected for more certainty:

> **"Collection"** means a collection of literary or artistic works, such as encyclopedias and anthologies, or performances, phonograms or broadcasts, or other works or subject matter other than works listed in Section 1(f) below, which, by reason of the selection and arrangement of their contents, constitute intellectual creations, in which the Work is included in its entirety in unmodified form along with one or more other contributions, each constituting separate and independent works in themselves, which together are assembled into a collective whole. A work that constitutes a Collection will not be considered an Adaptation (as defined above) for the purposes of this License.

The difference between a Collection and an Adaptation is that in the case of a Collection, "the Work is included in its entirety in unmodified form along with one or more other contributions, each constituting separate and independent works in themselves, which together are assembled into a collective whole". The difference is important because all the CC licenses authorize Collections, even the ND ones. A Collection was called a Collective Work in the versions prior to 3.0 by reference to the category of the US Copyright Act. The national qualification of Collective Work has consequences on the ownership of the work, which is vested according to many legislations on collective works not in the hands of the individual person who created the collection, but in those of the private or moral person who is responsible for directing the selection or the arrangement or funding the infrastructure (e.g.

---

[105] Indeed, there has been in 2006 a case of infringement of the synchronisation clause of a CC BY NC ND license by French national television; the grant to reproduce and publicly perform the work does not include the authorization to synchronize it on a documentary and as we will see further, CC music could not at that time be part of the catalogue managed by a collecting society which would have avoided the necessity of any prior request, televisions being used to declare titles afterwards. Dulong de Rosnay Melanie, "La musique de l'Onomatopeur reprise dans Envoyé Spécial sans son autorisation", *Creative Commons France blog*, 03-04-2006. http://fr.creativecommons.org/weblog/index.php?2006/04/03/45-la-musique-de-lonomatopeur-reprise-dans-envoy-spcial-sans-son-autorisation
The case has been settled out of court, the author received 500 euros, an amount equivalent to the royalty he would have received if he had had the option to be a member of the collecting society. This method is recommended to calculate damages in the considérant 19 of the Directive 2004/48/EC of the European Parliament and of the Council of 29 April 2004 on the enforcement of intellectual property rights (OJ L 157, 30.4.2004): "As an alternative, for example where it would be difficult to determine the amount of the actual prejudice suffered, the amount of the damages might be derived from elements such as the royalties or fees which would have been due if the infringer had requested authorisation to use the intellectual property right in question.

[106] http://creativecommons.org/weblog/entry/8213

the publisher, for instance the Wikimedia Foundation rather than the Wikipedian?), with respect to rights in the contributions which are retained by their original authors'.

d. Rights: Reproduce, Distribute and Publicly Perform

The rights granted by a CC license, notwithstanding when they apply to Collections and Adaptations, are expressed in 3 of the Definitions and include the rights to Reproduce, Distribute and Publicly Perform the Work:

> **"Distribute"** means to make available to the public the original and copies of the Work or Adaptation, as appropriate, through sale or other transfer of ownership.
> **"Publicly Perform"** means to perform public recitations of the Work and to communicate to the public those public recitations, by any means or process, including by wire or wireless means or public digital performances; to make available to the public Works in such a way that members of the public may access these Works from a place and at a place individually chosen by them; to perform the Work to the public by any means or process and the communication to the public of the performances of the Work, including by public digital performance; to broadcast and rebroadcast the Work by any means including signs, sounds or images.
> **"Reproduce"** means to make copies of the Work by any means including without limitation by sound or visual recordings and the right of fixation and reproducing fixations of the Work, including storage of a protected performance or phonogram in digital form or other electronic medium.

These rights' definitions are similar to some of the definitions of the Berne Convention article 11, the Rome Convention and the WPPT article 14, with some small differences, for instance "by any means of wireless diffusion" in Berne, "by any means" in the CC definition for Public Perform to broadcast and rebroadcast which is slighty broader and probably not problematic.

Publicly Perform includes "to make available to the public Works in such a way that members of the public may access these Works from a place and at a place individually chosen by them", but the right to Distribute "means to make available to the public the original and copies of the Work through sale or other transfer of ownership." Thus, because rental and lending are part of the right of making available to the public, but are not a transfer of ownership, it is unclear whether the rights of commercial rental and public lending are covered by the License Grant. This could be annoying because the grant intends to be as broad as possible and it should cover the commercial activity to rent videograms and the public lending by libraries of physical copies of CC-licensed works. Therefore, it would be recommended to include these two rights in the License Grant. But, these rights may lead to an "unwaivable right to equitable remuneration"[107] and be submitted to mandatory collective management provisions, without possibility for the Licensor to include them in the royalty-free grant.

[107] WCT article 7, WPPT article 9 (and also Rental Directive article 5).

e. The parties: the Licensor, the Original Author and You

The parties involved are the Licensor, the Original Author and You.

> **"Licensor"** means the individual, individuals, entity or entities that offer(s) the Work under the terms of this License.
> **"Original Author"** means, in the case of a literary or artistic work, the individual, individuals, entity or entities who created the Work or if no individual or entity can be identified, the publisher; and in addition (i) in the case of a performance the actors, singers, musicians, dancers, and other persons who act, sing, deliver, declaim, play in, interpret or otherwise perform literary or artistic works or expressions of folklore; (ii) in the case of a phonogram the producer being the person or legal entity who first fixes the sounds of a performance or other sounds; and, (iii) in the case of broadcasts, the organization that transmits the broadcast.
> **"You"** means an individual or entity exercising rights under this License who has not previously violated the terms of this License with respect to the Work, or who has received express permission from the Licensor to exercise rights under this License despite a previous violation.

We already noted in the description of the Attribution clause that it is not mandatory to identify the Licensor, the individual or entity that offers the Work. Because the Licensor offers the Work as indicated in the Definition, or grants the rights as indicated in the last sentence of the foreword, it can be assumed that the Licensor is the actual rightholder at the time the license is being issued, while the Original Author must actually intend to designate the original *rightholders* (in case rights have not been transferred, the Licensor and the Original Author will be the same person(s). These definitions could be clarified.

The definition for Original Author indeed encompasses:

- For artistic and literary works: the individual, individuals, entity or entities who created the Work, usually the *author*, or another entity (the film producer in the USA is recognized as an original author) but also the *publisher*, in case the author cannot be identified, perhaps in the case of orphan works, entities or publishers might be recognized as original rightholders in some jurisdictions, but this is not the case everywhere,
- For performances, the performers,
- For phonograms, the producer (again, neither the film producer nor the database producers are mentioned in case they are not recognized as entities who created the Work),
- For broadcasts, the broadcast organization.

Authors and other holders of rights related to copyright are to be identified in relation with the Attribution clause, requiring to provide the name of the original author. Indeed, Berne Convention article 15[108] states the principle of presumption of authorship: in the absence of proof to the contrary, the author is the person whose name appear on the work.

"You" designates the licensee, the person who has the authorization to exercise the rights granted by the License. But the Definition adds even more information. It anticipates on the Termination provision, by stating that a violation will end the License or, and this is not made explicit elsewhere in the License, that the Licensor may despite a previous violation grant express permission. It is not clear whether this targets violations which would have been performed, or exceptions made to the conditions that other licensees are deemed to respect. It

[108] As well as in the article 5 of the 2004 EC Directive on the enforcement of intellectual property rights.

is also not clear how this relates to the penultimate subclause of clause 8 stating that “This License constitutes the entire agreement between the parties with respect to the Work licensed here”.
Now that the main notions have been defined, we will review the clauses following their order of appearance in the Licenses.

ii. Fair Dealing Rights

The Fair Dealing Rights clause states that “Nothing in this License is intended to reduce, limit, or restrict any uses free from copyright or rights arising from limitations or exceptions that are provided for in connection with the copyright protection under copyright law or other applicable laws”.

In order to be truly international, this clause should be entitled Limitations and Exceptions because Fair Dealing is a national notion (UK, Canada, Australia). It could be also made clearer that limitations to related rights, and not only limitations to copyright, are not preempted by the License’s Restrictions and License Elements (for instance, that a performance can be parodied even if it is released under an ND license which reserves modifications).

iii. License Grant

The License grant is a worldwide, royalty-free, non-exclusive, perpetual license to exercise the rights described previously: Reproduce, Distribute and Publicly Perform, also in Collections, but in Adaptations only for licenses without the ND Element.

The royalty-free characteristic is limited by the clause related to collecting societies at the end of the Restrictions, the connection could be made clearer. This information, as well as the NC clause, could well fit here for all licenses, while it is currently the case only the non-NC ones, and the following section could be renamed for instance Notices and Credit. The clause related to technical measures could also be moved. Rights can be exercised in all media and formats, and technically necessary modifications are not considered to be Adaptations. These small modifications would improve the consistency of these complex texts which structure ends up being not logical.

The license intends to have the largest geographic and temporal scope possible: it lasts for the entire duration of copyright, but related rights or other applicable rights are not explicitly mentioned.

The license is non-exclusive, but it is not made explicit in the license that it is not compatible with exclusive licenses (such as underlined in the FAQs for rights assignments to collecting societies) or transfer of ownership and all exclusive rights through, for instance, a publication contract with an exclusivity clause. The information is not hidden and is obvious for the specialist, but not for the layperson who is often not aware of notions such as:

- The meaning of exclusivity,
- The prerogative of the original right holder to exercise her exclusive rights, and

- The impossibility to grant exclusive rights to a collecting society or a publisher when using a CC license.

Thus, a clarification could avoid licensors the risk of committing to incompatible agreements and be unable to comply with both at the same time.

iv. Restrictions

Many provisions contained in the section entitled Restrictions have already been studied: the BY and NC License Elements will thus not be analyzed again here.

The License states that the Work (but not the Collection apart from the Work itself), its copies and performances (videograms are not mentioned) can be made available to others only the terms of the License, which must be included under the form of a copy of the text or a link. Is this Notice requirement provision also applicable to the uses arising from limitations to exclusive rights? On the one hand, it should be the case in order to ensure CC-licensed works can be identified by the public and kept accessible as such, but on the other hand, requiring the notice to be kept intact can be interpreted as a restriction. It is indeed listed in the clause entitled Restriction while clause 2 states that nothing in this License is intended to restrict uses arising from limitations. Thus, it would be useful to clarify these two conflicting provisions, the URI or copy of the license must be included "with every copy of the work you distribute or publicly perform", and "nothing in the license is intended to restrict any uses free from copyright".

Besides, the mode of notification for performances is not provided. It often leads to questions by potential licensees working in analog or aural environments such as radios and exhibitions. By extension of the "reasonable manner" to implement the credit, the license can be indicated in a paper or online programme or on the wall of a venue, together with credit information.

No additional term or "effective technological measure" as named in the WCT and the WPPT (the 2 provisions could be paired to improve the readability) which would restrict the ability to exercise rights granted can be imposed by the licensee. Thus, any subsequent user should be able to access the work and exercise the rights granted by the license.

The collective management clause[109] is part of the restrictions for NC licenses and part of the license grant for non-NC licenses. It could be more closely related to the royalty-free provision that it amends. The goal was twofold.

First, announce to the Licensee that some uses may not be royalty-free:
- For non-waivable compulsory license scheme and
- For commercial uses (from Works which have been NC-licensed).

Second, prepare compatibility with collecting management schemes and authorize Licensors (the videogram producer being forgotten) to collect royalties:
- From non-waivable compulsory license schemes: for all licenses, even without the NC element,[110]

[109] See further developments *supra* in section 3. and 4.2.4.

- From waivable compulsory license schemes and voluntary license schemes: for commercial uses of works under NC licenses.

The final restriction is the moral rights clause. It has two components, one for all the licenses and one for licenses which authorize Adaptations, e.g. the licenses without the ND element. This clause informs the licensee that she should respect the moral right to integrity that the licensor may enjoy as part of applicable law. Authorized uses "must not distort, mutilate, modify or take other derogatory action in relation to the Work which would be prejudicial to the Original Author's honor or reputation", corresponding to the language of the article 6 *bis* of the Berne Convention stating that the author has the right to object to such actions. International law only foresees such a limit for authors, but not for other individuals or entities who are part of the CC definition for Original Author (the author, the publisher if the author cannot be identified, phonogram producers and broadcasters). One the one hand, the provision may impose more restrictions than the law as publishers usually do not enjoy moral rights. One the other hand, the provision may exclude some parties from its scope while they benefit from such a protection; moral rights may exist for non-authors in some jurisdictions, for instance for performers and film-makers in Australia, the latter being producers, directors and screenwriters, the film-maker producer being not mentioned in the CC definition for Original Author (she can be included if considered as a creator). Therefore, it is recommended to change the clause accordingly, and create distinguished definition (and a contact field to be filled by the Licensor when selecting her License) for Author and for the other Original Rightholders, in addition to Licensor who would be the current rightholder.

The second part of the clause, in order to waive some of the uncertainty on the possible conflicts between the right to allow the making of derivatives and the right to integrity, foresees that the licensor waives this right to the extend it is waivable ("to the fullest extent permitted by the applicable national law"). Indeed, in some countries like Japan, any adaptation could " be deemed to be a distortion, mutilation, modification or other derogatory action prejudicial to the Original Author's honor and reputation." However, regarding the situation of the countries where moral rights are not waivable, this clause has the drawback to imply that it might be actually impossible to authorize adaptations in advance after all, and therefore suggests a possible incompatibility of the non-ND licenses with moral rights. If it may bring certainty for jurisdictions like Japan, it sheds explicit light on a possible problem for jurisdictions in the other situation, in which "author's integrity may limit the extend to which one can freely license modification rights"[111] and might invalidate the license[112].

Besides this clause's two elements, other provisions of the licenses are related to the exercise of moral rights and reputation to a broader extend and could be placed nearer: obviously the attribution clause, but also the right to not be attributed upon request of any Licensor on Collections and Adaptations, and the non-endorsement clause stating that attribution should not imply a support by the Original Author, the Licensor or the Attribution Parties.

---

[110] It is to be noted that most societies do not allow their members to use a CC license, and those who introduced some sort of compatibility allow only the NC licenses. Thus, licensors are not in a position to join these societies and access the royalties. This clause is preparing the possibility.

[111] In the same sense, Välimäki Mikko, Hietanen Herkko, "The Challenges of Creative Commons Licensing", *Computer Law Review*, 6/2004 (Volume 5, Issue 6), pp. 172-177.

[112] More on moral rights in section 4.2.1.

v. Representations, Warranties and Disclaimer & vi. Limitation on Liability

We will now have a first look at clause 5 entitled Representations, Warranties and Disclaimer together with related clause 6 containing a Limitation on Liability.

As it is the case in most open source licenses[113], the Licensor offers the Work "as-is and makes no representations or warranties" including for product defects such as accuracy or merchantability, but also for noninfringement of third parties rights. The Licensor also disclaims liability for any damages arising out of the license or the use of the Work.

However, like the moral right waiver clause which was just discussed, CC licenses state in these two clauses that depending on the jurisdictions, these exclusions and limitations may be not applicable. Indeed, some consumer legislations forbid disclaiming certain warranties and some tort laws forbid misrepresentations[114]. Thus, these provisions will not be enforceable in all cases. Not all the CC licenses will contain such an exclusion and limitation. We will explain later in greater detail what are the arguments for both positions[115], and how the exclusion of representations and warranties of noninfringement and the limitation on liability for any damages relate to the security of the downstream chain and of the whole system in general[116].

vii. Termination

Clause 7 contains provisions related to the Termination of the license. If the licensee breaches any terms of the license, the license and the rights granted will terminate automatically. This affects only the License Grant (to Reproduce, Distribute and Publicly Perform the Work and Adaptations if applicable), and the Restrictions (requirements of copyright notice and Attribution, Non Commercial clause when applicable, waivers related to collecting societies and moral rights).

Otherwise, the license is perpetual for the duration of applicable copyright (and related rights even if they are not mentioned). However, the Licensor may stop distributing the Work, or distribute it under different terms, but these choices should not affect the license already granted or to be granted on existing copies of the Work which are available. This provision entitles licensors to make side deals. But the question of the right of withdrawal and the possibility to change one's mind is jeopardized by the nature of the Internet as old copies may still be available. Therefore, there might be at the same time copies of a work licensed under different conditions, the initial CC license which the Licensor had chosen, and the new terms, which can be another CC license or an all-rights-reserved policy. A licensor could not prevent usages based on the first license grant.

[113] Rosen Lawrence, *op cit*.
[114] See *supra* section 3.4 on the differences between jurisdictions.
[115] See *supra* section 3.2 on the previous versions of the licenses and the limited warranties clause in version 1.0.
[116] See *supra* section 4.2.3. on the effects for users of the disclaimer of warranty and liability.

viii. Miscellaneous

The eighth and final clause contains miscellaneous contractual provisions. When the licensee exercises the rights granted and distributes the Work or an Adaptation with a link to the License, the Licensor offers to the recipient a license to the Work on the same terms and conditions. As we will see in the coming section on the nature of the licenses, when Licensee B redistributes Licensor A's work to a third party recipient C, C gets a license from A, not from B, and this is also valid for Adaptations that B created based on A's original work.
The license contains a severability clause. As it has already been mentioned for warranties and liability, some provisions may be unenforceable in certain jurisdictions, and this should not affect the validity of the remaining provisions of the license.
A waiver of the terms of the license should be consented to in a written signed contract. This provision could be located closer to the provision allowing distributing the work under different conditions. It is slightly contradictory and then redundant with the penultimate subclause mentioning on the one hand that the license constitutes the entire agreement because another concluded at a later stage may exist elsewhere and on the other hand that the license may not be modified without the mutual written agreement of the Licensor and the Licensee.

The final subclause deals with international private law. It explains that rights and subject-matter were defined utilizing the terminology of the international conventions. Indeed we saw at the beginning of this analysis of the clauses that the definitions borrow, very largely but not entirely, from the definitions of the Berne Convention, the Rome Convention, the WIPO Copyright and Performances and Phonograms Treaties.
The core of the provision explains the rational of the porting by jurisdictions which will be analyzed in section 3.4: "These rights and subject matter take effect in the relevant jurisdiction in which the License terms are sought to be enforced according to the corresponding provisions of the implementation of those treaty provisions in the applicable national law."
The final provision clarifies some doubts which were raised in the definitions section: "If the standard suite of rights granted under applicable copyright law includes additional rights not granted under this License, such additional rights are deemed to be included in the License; this License is not intended to restrict the license of any rights under applicable law." This means that commercial rental and public lending rights, which are not mentioned in the scope of the rights granted, would be included. But this provision does not solve the question of subject-matter covered, namely whether first fixations of films and broadcasts and databases are covered.

2.3 The legal nature of the licenses

After having scrutinized the licenses' optional elements and main clauses and detected a few formal inconsistencies which would be possible to fix, we will now study the licenses as a whole and analyze their legal nature. We examined how the license clauses are compatible with copyright law. Now we will examine whether the licenses as tools are compatible with other area of private law, namely provisions governing contractual agreements or obligations, as well as more specifically provisions on unfair terms and consumer law regarding electronic and standard form contracts.

The licenses can be considered as licenses or as contracts depending on jurisdictions[117]. Beyond legal scholarship interest, it matters to identify the nature of the agreement in the scope of this study in order to identify possible incompatibilities with applicable law, assess risks and propose solutions to limit their consequences if they arise. Also, the legal qualification of the tools has an impact on the enforcement and the remedies options. It is important to know what are the possibilities in case of breach, otherwise the licenses would be worthless. It matters to find out first whether open licenses are licenses or contracts, because requirements for validity are different, they are much stronger for contracts; enforcement is also different, therefore applicable law (contract law or copyright law) and possible remedies for infringement (damages or injunction to enforce) will also vary as it will be further examined in section 4.1.

It also matters to verify that the agreement is valid and that consent between parties can be reached through such tools. These licenses intend to facilitate the use and the reuse of creative works, because permission is already granted and no additional transaction is required every time someone wants to use the work. Unlike traditional copyright agreements, from licenses of use to rights transfer contracts, neither the licensor nor the licensee sign any document to manifest their approval of the terms of an agreement allowing licensees to perform acts which would have otherwise infringed copyright. If the agreement is deemed invalid and consent has not been reached after all, permission will not deemed to have been granted. Licensors may not be able to request the enforcement of non-copyright infringement related conditions even if they apply to acts triggered by the exercise of a copyright-related right, and licensees might not be able to claim the exercise of rights beyond copyright law which is fully applicable by default, and thus reproduce the work freely.

Finally, it matters to identify a third specificity of the licenses, the share alike reciprocal effects and the transmission of obligations, therefore it should be analyzed if and how third parties may be bound by the conditions, otherwise the system would not be sustainable if the agreement enforceability stopped after the first round. Usually, obligations bind only the parties who consented to them and cannot be transmitted to third parties. But it is expected that the effect of the CC license will not stop after the first licensee and that the licensor will be able to enforce her conditions to subsequent users along the distribution and reuse chain to be built around the work to be redistributed, reused and modified.

---

[117] Guadamuz Andres "The license/contract dichotomy in open licenses: a comparative analysis", *University of La Verne Law Review* 30:2, 2009, p. 103.

In this section, we will therefore describe the legal nature of the CC licenses and interpret the possible consequences of the qualification of the Creative Commons texts, as well as their binding nature between parties and towards third parties. Their legal status will be studied according to validity, enforceability and termination arguments applied to the following parameters: the nature of these agreements (2.3.1), the formation of tacit consent based on behavior (2.3.2), the specificity of the transmission of rights and obligations (2.3.3). We will first explain the law applicable to contracts, licenses or obligations in some jurisdictions and then apply the theory to the CC licenses to analyze the nature of the legal deed and assess the licenses' validity, effect, and enforceability across jurisdictions.

Are there substantial differences between a license and a contract in terms of formation and enforcement? Are the necessary steps towards contract formation reached between the licensor and the licensee? What is the status and what are the consequences of non-negotiated unilateral agreements, end-user agreements, terms of use or standard forms agreements? Such questions are not only academic discussions, they are particularly relevant to assess the validity of the licenses, their binding nature and other legal effects and consequences for the compatibility of the system's expectations with the legal environment, for instance if there is a risk of breach of contract in addition to copyright infringement.

2.3.1 Unilateral permissions or contractual agreements?

What is a license? What is an open source or an open content license? What is the nature of a CC license? Several legal qualifications have been proposed for open source and CC licenses and the possible interpretations of the licenses will be discussed in this section. We already noticed that the machine-readable layer corresponds to a rights management measure and will concentrate here on the legal deed. Some scholars[118] studied the question of the nature of open source, open content and CC licenses, and several argumentations contemplate different solutions and teach diverging final conclusions: unilateral or standard contract, one-sided permission, non-contractual license, partial dedication to the public domain, limited abandonment, waiver, servitude, gift, promise…

Instead of detailing all the possible interpretations of the law and the literature, we will only review selected options to determine if the licenses are compatible with the law, if they fulfill

---

[118] Including Dusollier Séverine, "Sharing Access to Intellectual Property through Private Ordering", *Chicago-Kent Law Review*, 2007, p. 1391-1435; Guadamuz Andres "The license/contract dichotomy in open licenses: a comparative analysis", *University of La Verne Law Review* 30:2, 2009, pp. 101-116 and "Viral contracts or unenforceable documents? Contractual validity of copyleft licenses", *European Intellectual Property Review*, vol. 26 (8), 2004, pp. 331-339; Hietanen Herkko, "A License or a Contract, Analyzing the Nature of Creative Commons Licenses", NIR Nordiskt Immateriellt Rättsskydd (Nordic Intellectual Property Law Review), 2007/6, 76, p. 516-535; Guibault Lucie, van Daalen Ot, "Unravelling the myth around open source licences : An analysis from a Dutch and European law perspective", Information technology & law series 8, The Hague: T.M.C. Asser Press 2006; Pallas Loren Lydia, "Building a Reliable Semicommons of Creative Works: Enforcement of Creative Commons Licenses and Limited Abandonment of Copyright", *George Mason Law Review*, vol. 14, p. 271, 2007; Rosen Lawrence, *Open Source Licensing: Software Freedom and Intellectual Property Law*, Prentice Hall, 432 p., 2004; St Laurent Andrew, *Understanding Open Source and Free Software Licensing*, O'Reilly, 207 p., 2004; Shaffer Van Houweling Molly, "Cultural Environmentalism and the Constructed Commons", *Law and Contemporary Problems* 70, Spring 2007, p. 23-50; Välimäki Mikko, Hietanen Herkko, "The Challenges of Creative Commons Licensing", *Computer Law Review*, 6/2004 (Volume 5, Issue 6), pp. 172-177.

validity requirements, if and how they can be enforced between parties and beyond, we will focus on the main dichotomy between common law and civil law systems and possible qualifications of license or contract, to make sure that CC authorizations are valid permissions for the licensees and can be enforced by the licensor. The qualification has an impact on the different nature of claims remedies and damages available in case of breach of contract/license and/or copyright infringement.

We do not find a definite answer on the qualification of the CC licenses from the organization. On the one hand, the text of the licenses which foreword states "By exercising any rights to the Work provided here, You accept and agree to be bound by the terms of this license. To the extent this license may be considered to be a contract, the Licensor grants You the rights contained here in consideration of your acceptance of such terms and conditions." We only find certain hints, saying that the license might be interpreted as a contract[119], and the use of the words "acceptance" and "consideration" which are prerequisite to build a contract. On the other hand, it has been argued that the licenses are intended to be licenses, not contracts, as their name logically infers[120].

In order to understand the controversy, it is important to explain what is a license and what is a contract in both common and civil law, as they have different definitions and consequences in different legal systems.
A license is a unilateral act, a permission to do something which would otherwise not be permitted by law[121]. A driver's license is an example of unilateral permission granted by the state to an individual, there is no agreement or contract. A copyright license is a grant of a right which would otherwise belong to the exclusive rights of the right owner: without a license, exercising this right would be a copyright infringement.
A contract is a binding agreement between parties to do something creating obligations for both sides. It requires an offer and an acceptance in both civil and common law, which will be examined in section 2.3.2 about the consent. In addition, common law foresees a third factor to qualify as contract: the consideration, or "mutual obligation that is created by the agreement"[122]. In a unilateral contract, only the licensor makes a promise, while in a bilateral contract, both parties have obligations[123].

The main difference between a license and a contract is that a contract must meet material requirements to be formed: the offer and the acceptance, as well as the consideration in common law countries. In a license, the licensee does not have to be named[124]. If validity conditions were not met and CC texts could not qualify as contracts, they could still achieve something as non-contractual licenses and be enforced according to copyright law. This argument could satisfy American lawyers who may be afraid of the fragility of the loose structure of an open license (it does not identify the parties, there is no signature, no meeting between the parties) and that a judge wouldn't accept them as a valid contract based on the

[119] It is interesting to note that this mention only appeared at the version 3.0, maybe implying that the qualification was before that out of question for the headquarters.
[120] One example of lively discussion between Lawrence Lessig and CC affiliates from many jurisdictions on the qualification of license or contract is reported in Guadamuz Andres "The license/contract dichotomy in open licenses: a comparative analysis", *University of La Verne Law Review* 30:2, 2009, pp. 101-116.
[121] St Laurent Andrew, *Understanding Open Source and Free Software Licensing*, O'Reilly, 2004, 207 p. 4.
[122] Guibault, Van Daalen, p. 34.
[123] Rosen, p. 51.
[124] Hietanen, "A licence or a contract", p. 10.

lack of valuable consideration[125] as the licensor does not get a remuneration[126]. Other reasons provided to support the qualification as a license which would not be a contract are inherent to the US legal system: the difficulty of contractual disputes, the fact that contract law vary from state to state. But these statements are not convincing arguments, they reflect mere preferences and qualification is not a matter of personal choice or convenience. Besides, they were apparently limited to one country (the United States) and/or one school of thought (the Free Software Foundation), which case law is evolving: the Jacobsen dispute recognized the restrictions of the Artistic license to be of contractual nature. Even if the drafters of the GNU-GPL and of the CC licenses intended them to be licenses and not contracts, the qualification does not depend on their strategy. Anyway, in civil law countries and also according to many interpretations in common law jurisdictions, a contract is created by open licenses and therefore, an open license is a contract[127].

Finally, even if a license does not require consideration, which might be a convenient qualification if the requirement was not fulfilled in the US[128], there are arguments in the best interest of the CC system to avoid the qualification of mere license and seek the protection of the legal status of contract law.
First, a license is revocable[129] and can be terminated after 35 years according to US Copyright Act, and revocation raises uncertainty issues for the public if they are not sure the material will be permanently reusable. The text of the licenses itself says that the CC licenses cannot be revoked by the licensor, but only terminated in case of breach of the provisions. Thus, if a licensor revokes the license, it will not invalidate past usages, but what happens to licensees who would find copies and want to reuse them after the revocation without being aware of that fact?
Second, indeed without accepting a license, copying the work would be an infringement. But without contract law, it could be that some provisions of the CC licenses could not be enforced by the licensor[130]. Claims based on the rights granted (article 3) may be copyright infringement and protected as such, but the non-respect of provisions of the license restrictions (article 4) which are not related to copyright law, would be left without protection thought breach of contract or copyright infringement. If they were to be unenforceable, they would be worthless. However, this distinction between conditions within the scope of copyright and conditions outside the scope of copyright is fragile and the Jacobsen case decided the contrary. The conditions outside the scope of copyright suspected to need to rely on contract law apply to a work being reproduced, performed, distributed or modified, and these acts are copyright-related[131].

---

[125] We disagree with this fear that distributing a work under an open license would lack of consideration: the counterpart is free distribution, therefore promotion and fame, see argumentation on the absence of remuneration in the 2nd FAQ of CC France website at http://fr.creativecommons.org/FAQjuridiques.htm
[126] The GPL is a License, Not a Contract, Which is Why the Sky Isn't Falling. *Groklaw*, 2003. http://www.groklaw.net/article.php?story=20031214210634851
[127] St Laurent, p. 148, Rosen, p. 57, Guibault and Van Daalen, p. 34, Guadamuz (2009).
[128] Interestingly, a "Deed", the term chosen by the organization to name the summary even if CC claims it has no legal value, is enforceable without consideration and allows third-party beneficiary to enforce it, overcoming the privity issue: http://en.wikipedia.org/wiki/Deed (last visited 05-02-2010). The CC license does not fulfill the requirement of signature to be considered as a deed, but previously the requirement was a seal so evolution is possible.
[129] Pallas-Loren, p. 4 and 20.
[130] Dusollier Séverine, "Sharing Access to Intellectual Property Through Private Ordering", op cit, p. 1422.
[131] Shaffer Van Houweling Molly, "The New Servitudes", *Georgetown Law Journal*, Vol. 96, p. 885, 2008, draft available at http://ssrn.com/abstract=1028947, p. 52, note 282.

As a last remark, neither licensors nor licensees have an interest to deny the existence of a contract and start a lawsuit based on that ground: they usually need their conditions to be enforced and their licensed rights to be granted[132].

2.3.2 Consent to online non-negotiated texts

Now that we explained the substantial irrelevance of the debate to qualify the licenses as licenses or as contracts for the purpose of this study to ensure enforceability, we still need to demonstrate whether the licenses fulfill validity requirements. We will examine them according to laws which govern the validity of agreements.
We already approached the question of the formation of contract, requiring the manifestation of consent, the acceptance of an offer, as well as consideration in common law jurisdictions. We will therefore study how the licenses may build consent between the licensor and the licensee around the license grant and obligations.
We will consider the law governing general obligations, online and non-negotiable agreements, such as click-wrap, shrink-wrap, browse-wrap and standard forms and apply it to the CC licenses.
Verifying the compatibility of the licenses with both contract and consumer law is important to confirm their validity and their enforceability.

In civil law countries, contractual validity relies on formal elements such as manifestation of consent, the clarity of the notice and the information, the capacity of the parties, the legality and determination of the object of the contract.
Manifestation of consent, a condition of validity of contractual obligations, can be traditionally obtained when two parties shake hands, sign a document or click on a form as the law has extended the notion of consent and recognizes the validity of electronic contracts when the licensee is aware of the terms.

In Dutch law like in any civil law country, contracts are formed by an offer and an acceptance, they require an intention to produce legal effect, the intention being manifested by a declaration, or "the impression created by someone's apparent intention to produce juridical effects" (…), it may also be inferred from conduct. [133]

In French law, it can also be inferred from the fact that the recipient of the offer starts to execute the contract that it reveals her acceptation[134]. An offer in French law is the manifestation of will by which a person expresses to one or more, defined or undefined persons, the conclusion of a contract under certain conditions[135].

According to the Principles of European Private Law (an harmonization, codification and interpretation initiative by a group of scholars commissioned by the European Union), contractual enforceability is also granted to unilateral acts, to "any statement of agreement,

---

[132] Rosen, p. 66.
[133] Article 33 of Title 2 of Book 3 of the Dutch Civil Code, articles 3:35 and 3:37 (1), Guibault, Van Daalen, p. 40-41.
[134] Article 1985 of the French Civil Code, Dir. Michel Vivant, *Lamy Droit de l'Informatique et des réseaux*, par. 875.
[135] Dir. Gérard Cornu, *Vocabulaire Juridique Association Henri Capitant*, PUF Quadrige 4ème éd. 2003.

whether express or implied from conduct, which is intended to have legal effect as such", which would be "binding on the person giving it if it is intended to be legally binding without acceptance".[136]

These definitions of acceptance can be transposed to the CC licenses: the making available of the work by the licensor constitutes an offer, and the use of the work by the licensee (corresponding to actions granted by the license which would have otherwise constituted copyright infringement) is the manifestation of intention, the acceptance. Therefore, consent is expressed by behavior, even if the agreement is not simultaneous for both parties who will not meet – the licensor may well even never be aware that her licensed work found a licensee, someone exercising one or more of the rights offered by the license grant.

In American contract law, contracts require offer, acceptance and also consideration[137]. We already saw that according to some lawyers, consideration is not necessarily perfected by open licenses because no price is paid. But the free distribution and promotion of the work by others, otherwise a costly activity[138], as well as the Share Alike clause[139], are real and not illusory considerations.

Copyright contracts have very strict formal requirements under French law and if they are not met, the contract is deemed invalid and the rights not granted. Therefore, it should be checked if CC licenses would satisfy this formalism.[140] As they define precisely the extent (the rights granted at article 3), the duration (the duration of copyright), the location (worldwide) and the destination of the contract (the intention to contribute one's work to some sort of commons by authorizing some uses for free), we can conclude that the licenses meet the necessary formalism, which originally aimed at protecting authors against too broad transfers to publishers.

Other principles of private law intend to protect the licensee as a consumer in online and electronic distant or standard non-negotiable agreements against unfair terms and also impose requirements to the conclusion of the agreement, the acceptance step. We will now address the law governing agreements such as click-wrap, shrink-wrap, browse-wrap and standard forms also in common law and civil law selected jurisdictions to make sure that the tacit acceptance deduced by the use of the work is valid and binding or how the formal information process could be improved for more clarity and security.

In the US[141], online contract formation requires giving adequate notice of terms with three criteria: prominence, placement and clarity so that a customer will find and understand it easily, and express unambiguous assent. We will consider the situation of clickwrap,

---

[136] §I:101(2) and 103(2) in Study Group on a European Civil Code/Research Group on EC Private Law, von Bar Christian, Clive Eric (ed), *Principles, Definitions and Model Rules of European Private Law - Draft Common Frame of Reference (DCFR)*, Sellier, 2008, p. 183 cited by Guadamuz, "The license/contract dichotomy", p. 111.
[137] Rosen, p. 59-65.
[138] See our argumentation on the absence of remuneration in the 2nd FAQ of CC France website at http://fr.creativecommons.org/FAQjuridiques.htm
[139] Guadamuz, "The license/contract dichotomy", p. 108.
[140] More details on the application of article L. 131-3 of the French Intellectual Property Code to the CC licenses in the 2nd FAQ of CC France website at http://fr.creativecommons.org/FAQjuridiques.htm
[141] The following analysis borrows from Kennedy Charles H., Making Enforceable Online Contracts, *Computer law review international*, 2009, n°2, pp. 38-44.

shrinkwrap and browsewrap agreements.
Clickwrap scenario provides strong evidence that the customer by clicking on a button asserting "I agree" read the proposed contract. Some CC public domain tools require clicking on a button to express the agreement, but the standard licensing suite does not offer this step.
Shrinkwrap contracts must also comply with these requirements on effective notices. The use of the product is binding if the user had the opportunity to review the notice, according to ProCD v. Zeidenberg case[142] or he could have returned the product. Inconsistency in naming the terms and confusing documentation should be avoided: it must be clear that the terms are a binding contract[143]. Therefore, there is a small concern due to the non-binding nature of the Human Deed and the risk of confusion with the Legal Code.
In browsewrap contracts however, the user does not exercise such an assertive action expressing her assent. The Specht v. Netscape Communications[144] case reveals that "the mild request 'please review',… reads as a mere invitation, not as a condition. The language does not indicate that a user must agree to the license terms before downloading and using the software… A reference to the license terms on a submerged screen is not sufficient to place consumers on inquiry or constructive notice of those terms". Only a "Download" rather than an "I agree" button was deemed insufficient. But the software was monitoring online activities while a CC license does not have such negative hidden terms as it allows using a work which would otherwise be submitted to exclusive rights. However, the disclaimer of representation is an inconvenient of the product and clear notice that the work may be infringing others' rights requires reading the Legal Deed.
We should also note that the language indicating the terms corresponding to the Notice of the CC license must be placed by the Licensor on her website. Therefore, the burden on explaining precisely with a clear sentence in the License Notice that the logo corresponds to the licensing terms relies on the Licensor who downloads a license from the user interface.
The opportunity to review the terms is also facilitated by a clear graphical presentation and language. We will come back to these arguments in section 5 to support the use of plain language instead of legalese jargon, and to advocate for the development of more tutorials to help licensors to accompany the making available of their works under a CC license by a well-designed interface and a clear Notice language to indicate the hyperlink to the license. Even in the case of the qualification as a license and not as a contract and therefore no obligation to respect these validity requirements, more clarity could only benefit the system.

In both US and European systems, the recipient must also be able to store and reproduce the terms, which is the case with the CC licenses which are easily and permanently accessible online. But European case law has been less strict: a German Court recognized that terms of the GPL were part of the contract because a reference was made on a webpage[145], and a Dutch Court[146] decided held the acceptance of the CC terms valid because the infringer as a professional should have checked the terms. The conditions apply even if the other party hasn't read them. In case of doubt, the magazine should have contacted the author, as in a regular transaction in a classic all-right-reserved copyright environment.

---

[142] ProCD, Inc. v. Zeidenberg, 86 F.3d 1447 (7th Cir., 1996).
[143] Kaufman v. American Express Travel Related Services, Inc., United States District Court, No. 07 C 1707, 2008 WL 687224 (March 7, 2008).
[144] Specht v. Netscape Communications, 150 F. Supp. 2d 585 (S.D.N.Y.2001).
[145] Landgerichts München I, 19/05/2004 No 21 O 6123/04. http://www.jbb.de/urteil_lg_muenchen_gpl.pdf
[146] District Court of Amsterdam, Adam Curry v. Audax Publishing B.V., Case 334492/KG 06-176 SR, 9/03/2006; European Copyright and Design Reports, Sweet and Maxwell, Westlaw, septembre 2006, Curry v Audax Publishing BV, 2006 WL 2584400, [2006] E.C.D.R. 22 (RB (Amsterdam), Mar 09, 2006) (NO. 406921).

But this last decision did not involve a consumer. Indeed, Dutch law makes a distinction between professionals and consumers who may download a work only because it is accessible for free, "without realising that a license governs its use".[147]
Also, these decisions were related to simple cases of infringement of rights of the author by the first user, not involving non-copyright related conditions nor a chain of derivatives and subsequent users.
For Séverine Dussolier, "the mere fact of using the licensed object, modifying it, or distributing it does not mean that the user is aware of all the terms and conditions and has accepted them"[148]. For Lucie Guibault, "a user would be bound to the license terms as a result of his actions only if he actually accepted the legal consequences of his actions, and accomplished these actions with the specific intention to be bound by the license." The use of a hyperlink to indicate conditions can be compliant to Dutch contract law if the link is in a visible place, thus probably not by posting such a link at the bottom of the homepage[149].

Clickwrap methods offer indeed safer legal evidence of consent, but in practice nothing proves that the user read the terms even if she had the opportunity to do so as she may click on "I accept" without having read them. Even if the CC licenses are visible enough to be binding, it could be useful to further develop the acceptation interface, which is constituted by the download interface for the licensor and the notice text for the licensee.

The European Directive on Electronic Commerce[150] and its Dutch implementation[151] requires providing clear, comprehensible and unambiguous information[152] as well as the technical steps to follow to conclude a contract.
These requirements do not fit to the architecture of the CC project, which does not keep track of generated licenses or licensed works, unlike to the expectations of many licensors as shown by the large amount of questions inquiring whether CC will store information related to the licenses applied to works.

The French transposition of the European Directive on Electronic Commerce, the law on "confidence for digital economy"[153], requests a double signature[154] to translate the consent of

---

[147] Guibault, Van Daalen, p. 43.
[148] Dusollier, Sharing, p. 1424.
[149] Guibault, Van Daalen, p. 43 and 47.
[150] Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market, 17/07/2000, OJCE L. 178/1.
[151] Aanpassingswet richtlijn inzake elektronische handel, Stb. 2004, No. 210 and Dutch Civil Code article 6:227b(1). The following developments are borrowed from Guibault, Van Daalen p. 41-51.
[152] Ibidem p. 41.
[153] Loi n° 2004-575 du 21 juin 2004 pour la confiance dans l'économie numérique, JORF du 22 juin 2004, p. 11168.
[154] The "double-click" process ensures that people who buy or download a product learn the use condition AND accept them through clicking on a button "I read the conditions and accept them" and a new window must appear "you are downloading this under these conditions, you recognize having read and accepted it" which must be follwed by a button "I accept". This procedure is compulsory for persons acting professionally as licensors even if nothing is sold. An implementation procedure is to allow the download only after the user has displayed the license (not the Notice button) and expressed her agreement through a separated mouse click. The beneficiary of the "offer" must have the possibility to verify the "order" details and price and correct possible mistakes, before confirming the offer to express his acceptance. The issuer of the offer must acknowledge receipt of the order. Professional offers must decribe the steps to conclude the electronic contract and technical means to allow the

a consumer and the constitution of a contract with binding obligations, in order to make sure the consumer is aware of her agreement. Without this formalism around the acceptance of the condition of use, online contracts are not valid, even if nothing is being sold, and it also applies to the provision of information through download and browsing. Therefore, it is questionable whether the method to become a CC licensor should implement a double-click mechanism. However, mere access to the work, or use of the work following a limitation to exclusive rights, do not require either a CC license permission as these acts are outside copyright law regulation.
It could be that neither CC providing legal documents nor the Licensor offering a Work under a CC license are in the scope of this law because there is no order or individual request between CC offering licenses and the potential user, the Licensor who can use at will the "choose license" interface without pasting the code next to her work, and because unlike to a downloaded software, it is not because a user browses or downloads a CC work that she will exercise one of the additional freedoms, and make more than a personal or fair use which does not deserve any licensing agreement. If the "offeror" who should respect this double click provision should be one of the two CC license parties, it should be first identified whether it is the Licensor or the Licensee who performs the "characteristic service provision", criteria to identify who is the weak party, usually the consumer, to be protected: the Licensor who offers her work for free, or the licensee who will be able to exercise certain acts on the work only if she fulfills certain conditions.
Because it does not seem a good idea to burden CC interface with additional text before download a license or browsing a licensed work, it could be a solution to explain in the FAQ that licensors may want to insert additional information or an interface in their websites proposing CC works.

The Directive on Electronic Commerce and already the Directive on Distance Contracts[155] require the service to provide identification information such as a name and a physical or electronic mail address. This requirement may be implemented by informed parties, but it is not enabled by the CC interface and it has already been suggested in the previous section to provide a contact for the Licensor. This could also be added in the FAQs.

The last step in European law to pass in order to be valid is consumer legislation against

---

beneficiary, before the conclusion of the contract, to identify possible mistakes made in the data typing, correct them if relevant, and confirm to express his acceptation. This procedure has been enforced by a free software license, CECILL, developed by three institutions of french public research, informing on the license website that offering softwares under a CECILL license is conditioned by the reading of the license and its approval to avoid possible liability and respect consumer legislation. The website provides guidelines for licensors to implement on their websites to distribute software under a CECILL license and respect the formalism of the electronic commerce legislation:
The free software should not be downloaded before all these steps are fulfilled by the licensee who accepts the offer:
- The license must be readable on the website proposing the software download,
- The person who wants to download the software must before this click on a button "I accept the terms of the CECILL license that I read",
- After this click and before effective download, the user must see a new window with a warning "you are about to download a software under a CECILL license that you have read and accepted",
- Last window must be validated by a click "I accept" which closes the contractualisation process and valids licensee consent. (Source: our translation from http://www.cecill.info/mode-emploi.fr.html)

[155] Directive 97/7/EC of the European Parliament and of the Council of 20 May 1997 on the protection of consumers in respect of distance contracts, OJCE L 144, 04/06/1997, p. 19-27, article 4.

unfair contractual terms[156]. The exoneration of liability clause[157] and very detailed attribution requirements could be declared invalid.

[156] Council Directive 93/13/EEC of 5 April 1993 on unfair terms in consumer contracts, OJCE L 095, 21/04/1993, p. 29-34.
[157] This point will be further analyzed in section 4.

2.3.3 Transmission of obligations to third parties

After discussing the formal requirements to ensure the offer by the licensor and the acceptance by the licensee are valid regarding contract and consumer law, we will now consider the effects of a CC license on subsequent derivative works and on third parties reusing these works, and discuss the enforceability of the Share Alike clause.

We will first explain how the CC licenses intend to bind subsequent users after an initial licensor/licensee direct relationship, and how the system builds a distribution and licensing chain of generations of unmodified and/or modified works. We will analyze the relation between parties in a scenario involving more than two initial parties. Does the user at the end of a chain of derivatives get a license from each of the successive contributors, and also from the licensor of the Original Work or only from the immediate predecessor?

It matters that the CC license is not only enforceable against the immediate licensee, but also against third parties subsequent users. Otherwise, if an author A releases a work under a BY-NC-SA license and an author B modifies it, C could, for instance, make a commercial use of the derivative because she has no contractual relationship with A.

We will study how the Share Alike clause might bind subsequent users according to the concept of passing obligations to third parties, which is called privity in English and American contract law. It is a general principle in civil contract law that only parties to an agreement are bound by it, in order to protect parties from being subjected to burdens they would not be aware of. Therefore, the transmission of obligation to third parties must be further studied in common and civil law jurisdictions in order to understand if and how terms can follow the work and bind subsequent users.

We will finally consider the sublicensing option, which has not been chosen by CC as a licensee is not allowed to sublicense the Work.

Contract-as-products accompany software products and works available online. The specificity of open licenses is that obligations will follow the product when reused by third parties. Open licenses are qualified as "viral contracts"[158], "contracts whose obligations purport to 'run' to successor of immediate parties" because they bind subsequent users, the Share Alike provision requiring derivatives to be licensed under the same terms. Also, each licensee must include a copy of the license or a link when distributing the Work.

CC licensing facilitates the redistribution of works in an unmodified or modified version. Therefore, a cascade of rights, obligations and responsibilities circulates together with the work all along its lifecycle. A long chain of parties who do not have a direct link with the original licensor can thus be constituted. The licenses are expected to bind downstream parties, otherwise licensors may be reluctant to offer their works if their conditions are not respected after the first copy or modification into a derivative work.

---

[158] Radin Margaret Jane, "Humans, Computers, and Binding Commitment", *Indiana Law Journal*, vol. 75, p. 1125-1161, 2000.

The definition of the legal relation between the licensor and the subsequent licensees will impact on the possibility for the initial licensor A to sue a second-range licensee C, or the second-range licensee C to sue the first licensee and second licensor B if C committed an infringement of A's rights without knowing it because the first licensee B did not properly respect the terms of the license granted by A.

A cascade of infringement may be transmitted to subsequent authors of derivatives who would ignore that the first derivative did not for instance properly acknowledge the original author[159]. The disclaimer of warranties gives few legal security to licensees and does not incentivize users to rely on the usability of CC-licensed works. Each new action performed on the work implies the formation of a new relation between the parties, A and B and then B and C as well as A and C. "There must be an unbroken chain of privity of contract between each successive user of the content"[160].

Let us now examine how the CC licenses foresee to implement the principle of privity to pass obligation from A to a subsequent licensee C.

We already discussed the confusion between Original Author, original right holders and Licensor in the Definitions section. Let us assume in this section for the purpose of distinguishing problems that the Licensor A is the only original author and sole initial right holder.
The Licensee B is the person who will reproduce the Work, distribute it in a Collection, or create and distribute an Adaptation
According to article 4.a., the Licensee B may not sublicense the Work, and according to 8.a. and 8.b, when the Licensee B distributes the Work or a Collection or an Adaptation, the third recipient C enters into a relation with the Licensor A.
In the case of article 8.b, Licensee B made an Adaptation Y of the Original Work X licensed under a Share Alike license. C wants to make another Adaptation Z. Therefore, C will be the Licensee of B for Work Y and the Licensee of A for Work X.

Will C be aware when she reuses Work Y that she enters in a relation not only with Licensor B but also with Licensor A? It can get complicated if B did not properly acknowledge A, or if A asked her name to be removed, or if B did not explain properly the modifications between Work X and Work Y.

Let us now take the case of Work X offered by Licensor A. Licensee B incorporates Work X without making an Adaptation of it into a Collection XYZ. Collection XYZ, on the one hand, does not have to be distributed under a Share Alike clause, but on the other hand, when Licensee B distributes the Collection XYZ, it seems that:

- Licensee B cannot sublicense Work X to C so licensee C will not have a relation with A through B but directly with A
- By the virtue of clause 8.a, Licensor A offers to recipient C the Work X and the Collection XYZ under a Share Alike license.

---

[159] Elkin-Koren, p. 418.
[160] Merges Robert, "A New Dynamism in the Public Domain", *The University of Chicago Law Review*, Vol. 71, p. 199, 2004.

- There is no relation between B and C and B did not have to release the Collection XYZ apart from Work X under a Share Alike license.

It becomes very complicated, especially after more than three parties, collections and adaptations, and all the more if the identification and contact of the parties are not available.

Now that we examined how the CC licenses foresee to implement the principle of privity to pass obligation from A to a subsequent licensee C, let us see if and how a contract may be automatically concluded every time the work is distributed, e.g., between A and C, and if therefore A can sue C if C does not respect the Share Alike clause.

In English common law, the principle of privity prevents to pass burdens to third parties but makes it less difficult to pass benefits[161]. In civil law, the Share Alike clause is questioned by the general principle of the relative effect of contracts and of the difficulty to bind third parties. Solutions might be found in clauses related to the relative effect of contracts in the case of positive rights created to the benefit of the third person[162]. But some doubt that the Share Alike clause succeeds into creating contractual privity between the licensor and each of the licensees[163], which brings back to the question whether A could sue C for copyright infringement or for breach of contract in case of non-respect of the Share Alike clause.

Despite doctrinal difficulties to justify the validity of relative effect of the contract, enforcement cases revealed the validity of several licenses copyleft clauses and not only in simple case with only one direct relationship between two parties.

In Jacobsen v. Katzer[164], the Court decided that the attribution conditions of the Artistic License on the use of the modifications are contractual obligations. A French Court decided in 2009 that the Licensor was bound by the GNU-GPL to deliver the source code to the Licensee and to include a notice to the license[165]. It is remarkable in this case that Licensee C won over Licensor B who had removed notice and attribution of Licensor A without Licensor A being involved in the lawsuit.

But two options are available to guarantee the enforceability of the licenses terms along the distribution and modification chain: the Share Alike clause and sublicensing. Sub-licensing is actually excluded by the licenses, which makes it impossible to have a direct relationship between each successive parties and then have B endorse some responsibility towards C and allow C to sue B if A sues C while B committed the infringement. Maybe the question of sub-

---

[161] Guadamuz, *Viral contracts*, p. 336-337.

[162] Article 6:253 of the Dutch Civil Code: "A contract creates the right in favour of a third person to claim a prestation from one of the parties or to invoke the contract in another manner against one of them, if the contract contains a stipulation to that effect and if the third person accepts it"
Article 1121 of French Civil Code permits to waive the consent requirement: "One may likewise stipulate for the benefit of a third party, where it is the condition of a stipulation which one makes for oneself or of a gift which one makes to another. He who made that stipulation may no longer revoke it, where the third party declares that he wishes to take advantage of it."

[163] Guibault, Van Daalen, p. 53-56.

[164] United States Court of Appeals for the Federal Circuit, No. 2008-1001, Robert Jacobsen v. Matthew Katzer, 13-08-2009.

[165] Cour d'Appel de Paris, No 04/24298, Edu4 v. AFPA, 16-09-2009. Summary in English available at http://fsffrance.org/news/article2009-09-22.en.html

licensing should be re-considered[166], so that B could license A's work, but it is a tricky issue because rights are not transferred or exclusively assigned though the license. Currently, the only way for a licensee to become a licensor is to create a Derivative Work.

[166] Guibault considers that option for the GNU-GPL, see Guibault, Van Daalen, p. 54-55.

3. Sources of potential incompatibility

Now that we studied the licenses clauses and their possible incompatibilities with copyright and contract law, we will examine possible incompatibilities within the system. Internal incompatibilities will be tracked down among the licenses' different versions, options, jurisdictions and with compatible licenses. Some of them are visible incompatibilities, for instance it is well known that not all option combinations are compatible: it is not possible to remix works licensed under incompatible options. But some incompatibilities or inconsistencies are not easily ascertainable to the user.

Trying to cover the spectrum of rights between full copyright and the public domain raises another issue: paradoxically, all the licenses do not support the remix culture based on combination, collage and reuse. The option reserving the right to make derivative works makes it impossible to adapt works. The multiplicity of options threatens interoperability, as works licensed under different Creative Commons cannot always be mixed to create a third work. The benefits of the system are therefore limited, because despite an apparent ease of use, the internal incompatibility often reduces the possibilities to the right of sharing the verbatim work for non-commercial purposes, without possibility to adapt it or to distribute it in commercial situations without further negotiation, just as in the traditional copyright system. The pool of works under a Creative Commons licenses is thus partly sterile, because most of the works cannot be recombined together to create derivative works without requiring additional permission.

This chapter will describe the differences between licenses which may cause incompatibilities and hinder the use of the works including the ability to remix them together. Two sources of differences are visible from the license chooser (formats and options) but actually five sources of differences between the licenses may raise incompatibilities issues:

- The licenses formats, the machine-readable code, the human-readable common deed and the legal code (formats: section 3.1),
- The licenses different options and combinations: BY, BY-SA, BY-NC, BY-ND, BY-NC-SA, BY-NC-ND (options: section 3.2),
- The licenses successive versions: 1.0, 2.0, 2.5, 3.0 (incremental versions: section 3.3),
- The differences between the licenses adaptations to various jurisdictions, the porting process has been engaged for the six combinations and at least one version for over 50 countries or jurisdictions (jurisdiction versions: section 3.4),
- The differences with other similar licenses which have the same purpose but use a different language and may become compatible with the BY-SA (other open content licenses: section 3.5).

These five sources of identified and unidentified incompatibility will be presented by order of level of visibility and difficulty they may raise.

The differences between the formats (section 3.1) and the incremental versions (section 3.3) as well as the differences between the options and the resulting incompatibilities between the

combinations (3.2) will be described systematically. Differences between formats and option combinations are visible to the user. They are not hidden in the texts of the jurisdictions legal deeds or previous incremental versions which require the user to both be aware of their existence and to look for them on the website by generating an other license or by modifying the license URL. They are accessible in plain English on the Creative Commons website, and resulting incompatibilities are easily identifiable.

However, the differences justified by the adaptation to local legislations (section 3.4) are less visible and may raise more complex issues. Some incompatibilities are hidden by the fact that licenses carrying the same license elements may cover slightly different rights and subject-matters after such rights have been defined according to different national laws.

Creative Commons jurisdictions licenses are deemed to be equivalent by virtue of the Share Alike compatibility clause[167], while their substance may diverge widely. Some of the international licenses provide a re-translation into English on the Creative Commons website, but it is difficult to assess the impact of these differences without a deep comparative legal knowledge. It is questionable whether jurisdictions licenses which have been adapted to national law are fully compatible among each other, for instance as some, but not all of them, include related rights or database rights. We will not study neither all the clauses nor all the jurisdictions for all the international versions, but select a number of representative points and countries.

Finally, the fifth source of potential incompatibility also involves licenses which are intended to be declared compatible (section 3.5) in the same vein as the international texts among each other. The Share Alike clause provides not only that international licenses are compatible, but also that licenses outside the Creative Commons system may be declared compatible, thus also allowing a re-licensing of derivatives under these licenses. The process has not been finalized and none of the licenses which may be seen as natural candidates given the similarity of their goals have been declared compatible yet. But it is a matter of time and political decision before some licenses are declared compatible and issues need therefore to be analyzed. It has been underlined since the birth of the licenses that paths should be found to facilitate the reuse of works licensed under a Creative Commons Attribution Share Alike license, a Free Art License, and a GNU Free Documentation (GFDL) license among other licenses. Until they are declared compatible, it will be impossible to synchronize for instance a CC BY-SA music track on a GFDL text-to-speech version of a text without asking permission to initial authors.

Despite the youth of the movement, there have already been three revisions of the licenses, thus four incremental versions which have been released in less than five years. The high number of available licenses' incremental versions has been caused by the need to fix the initial influence of US law and to solve some other individual problems. The two first versions of the licenses had indeed been written in reference to United States copyright law definitions. It is only with the fourth version 3.0 that the legal code generated by the headquarters became truly "generic" or "unported" by referring to international copyright law. Nevertheless, the internationalization of the licenses started from the initial version 1.0, and

[167] The Share Alike clause provides that the derivative of a work licensed under a Share Alike license may be licensed under the same license, but also under an international license with the same optional elements or a license which will have been recognized compatible.

over 50 jurisdictions already translated the texts and/or adapted their provisions to their national legislation. If all the countries had adapted all the versions, which is far from being the case, there would be about 50 countries per 4 versions, thus 200 sets of 11 option combinations, and then 6 option combinations: up to 1200 licenses in theory (probably around the half in reality, as most jurisdictions haven't ported all the versions).

Proliferation is endangering the sustainability of a movement intending to facilitate reuse, not to prevent it or to hide problems. As introduced in the previous chapter, two main critiques arise therefore from the licenses diversity for both licensees and licensors:
- The risk of missing one of the most preeminent opportunity and objective of the organization and of impairing the movement generativity if free culture can't even be applied within the system as most resources can actually not be recombined and remixed together,
- The risk of ideological fuzziness, in connection with high information costs to choose a suitable license among available optional elements.
Unforeseen legal consequences can be added to the list of risks, in the case of international and external licensed recognized as compatible but which contain substantial differences.

3.1 Incompatibility between different formats

The licenses exist in 3 formats readable by machines, humans and lawyers. The average user will only browse the logo, which displays the options and a link to the license incremental version. More experienced users, and that is the objective of the layers, will click on the logo and actually read the Common Deed. What are the differences between the information provided in all the formats? Not all users will click on the link at the bottom of the Common Deed to access the Legal Code. In section 4.1.2, we will come back to the impact of the existence of the three layers and their differences on contract formation and consent as the Commons Deed declares it is not binding. The Commons Deed is more accessible and there are more chances that people read at least some summarized clauses compared to other licensing schemes which only have a long and hard to read Legal Code. But this handy feature is irrelevant as a legal requirement to appreciate the consent. The fact is that it does not contain all the information and jeopardizes the informed assent of the licensee. It contains the summary of only selected clauses while many provisions are not mentioned.

The core grant in the human-readable deed states:

You are free:


to Share: to copy, distribute and transmit the work (in all the licenses)


to Remix: to adapt the work (in the non-ND licenses)

These two logos illustrate the right to reproduce, perform and distribute including in Adaptations. They could be used in other portions of the interface to express the positive grant

of the license.

The conditions are summarized next to the usual logos of the license elements:

Attribution — You must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work).
Share Alike — If you alter, transform, or build upon this work, you may distribute the resulting work only under the same, similar or a compatible license.
Noncommercial — You may not use this work for commercial purposes.
No Derivative Works — You may not alter, transform, or build upon this work.

Not all main clauses are summarized, only the following are included:

Waiver — Any of the above conditions can be **waived** if you get permission from the copyright holder.
Public Domain — Where the work or any of its elements is in the **public domain** under applicable law, that status is in no way affected by the license.
Other Rights — In no way are any of the following rights affected by the license:
Your fair dealing or **fair use** rights, or other applicable copyright exceptions and limitations;
The author's **moral** rights;
Rights other persons may have either in the work itself or in how the work is used, such as **publicity** or privacy rights.
Notice — For any reuse or distribution, you must make clear to others the license terms of this work. The best way to do this is with a link to this web page.

The waiver can be misleading. For instance Attribution is listed as a Condition while it cannot be waived in many countries with strong moral rights. Many of the main clauses are not summarized here, for instance the definition of work or of collection, the collecting societies clause, the disclaimer of warranties and representation, the limitation of liability, the termination clause, etc. Therefore, it could be possible that a licensee is not aware of important limitations such as the absence of representation or the fact that all uses will not necessarily be free as royalties might be collected by collective societies. And it is contractually more important to pay attention to possible approximations and omissions in the Commons Deed which does not represent fairly and accurately the binding information which is contained in the Legal Code.

The main clauses are summarized as follows on a webpage entitled “baseline rights” which is not prominently displayed, but seems highly relevant for our purpose of identifying clearly most of the rights and conditions for both parties without hiding too much information because one format is shorter than another[168]:

“All Creative Commons licenses have many important features in common.
Every license will help you

- retain your copyright
- announce that other people’s fair use, first sale, and free expression rights are not affected by the license.

Every license requires licensees

- to get your permission to do any of the things you choose to restrict - e.g., make a commercial use, create a derivative work;
- to keep any copyright notice intact on all copies of your work;
- to link to your license from copies of the work;
- not to alter the terms of the license
- not to use technology to restrict other licensees’ lawful uses of the work

[168] http://wiki.creativecommons.org/Baseline_Rights

Every license allows licensees, provided they live up to your conditions,

- to copy the work
- to distribute it
- to display or perform it publicly
- to make digital public performances of it (e.g., webcasting)
- to shift the work into another format as a verbatim copy

Every license

- applies worldwide
- lasts for the duration of the work's copyright
- is not revocable"

This summary differs substantially from the Commons Deed language, first because it is addressed to the Licensor while the Commons Deed targets the Licensee, but also because it focuses on the core clauses, while the Commons Deed focuses on the License Elements, in addition to a few more references to other clauses which have been added under the title "With the understanding that" after revisions[169] discussed with the users' and international affiliates communities (mostly that fair use, moral rights and other rights such as publicity rights are not affected).

The previous versions of the Commons Deed are not available anymore from the CC interface. The Internet Archive Wayback Machine[170] gives interesting results when searching for previous versions. For instance, http://creativecommons.org/licenses/by/1.0/ on February 1st 2004 was mentioning that the grant included the right "to make commercial use of the work". It is cognitively useful to also display the contrary of NC and the contrary of ND (commercial uses and derivatives allowed). Even if it seems tricky to change the licenses titles, a more coherent naming policy could be helpful, as the non-ND feature is currently not noticeable. As underlined earlier, the License Elements are more visible than the core clauses. We would recommend displaying both the non-NC and the non-ND rights in the relevant licenses combination for more clarity and thus indicate also rights which are not License elements, instead of featuring only License Elements which restrict from the positive grant.

[169] The revisions of the Commons Deed are not exactly synchronized with the versioning of the Legal Code.
[170] http://www.archive.org/web/web.php

3.2 Incompatibility between different versions

This section focuses on selected legal discrepancies reflecting debates and modifications between the licenses successive versioning. Policy debate and legal discussion took place among users and international affiliates communities at the occasion of each of the licenses versioning, on the mailing lists and for the last versioning during meetings which involved the international community.

Only the last version is available from the "Choose your license" interface and can be obtained from the Creative Commons website. However, previous versions are used on the web and available on numerous websites. Indeed, not all licensors use the interface to generate a license, it is possible to copy the logo from another website and thus not necessarily use the latest available version. Nevertheless, common deeds from previous versions contain a link to the newest version with the following mention informing licensors:

> A **new version** of this license is available. You should use it for new works, and you may want to relicense existing works under it. No works are *automatically* put under the new license, however.

As we will see in the subsequent section 3.3 presenting the potential incompatibilities between the licenses of different jurisdictions, not all jurisdictions are at the same porting stage or ported all the licenses. For instance, all the four versions are available in the Netherlands jurisdiction, while only version 2.0 has been ported in other jurisdictions.

It will be analyzed whether the differences between the successive versions create incompatibilities between licenses carrying the same optional elements, based on the list of the differences, intended to be improvements, at each of the versioning, as presented on Creative Commons blog.

3.2.1 From 1.0 to 2.0 in May 2004[171]

a) Attribution becomes standard

Attribution was an optional element in version 1.0, leading to 11 different licenses in combination with the other optional elements (Non Commercial, Non Derivative, Share Alike), the 6 current licenses and 5 additional which did not contain the Attribution element. As it had been observed that up to 97-98% of the users were selecting the Attribution element on the license chooser interface, Creative Commons organization decided that Attribution would not be optional anymore. This would contribute to drastically reducing the number of available licenses from 11 to 6, and users would also have one less question to answer on the license selection interface. That option is standard in many copyright legislations, but not in US copyright law, or in a very limited manner for visual artists.

[171] http://creativecommons.org/weblog/entry/4216

b) Share Alike compatibility with future and international versions

The version 1.0 licenses required derivatives to be published under the exact same license only. Version 2.0 states that derivatives may be re-licensed under one of three types of licenses: (1) the exact same license as the original work; (2) a later version of the same license as the original work; (3) an iCommons (which has been renamed CCi in the meanwhile license with the same license elements as the original work).

Thus, a work under BY SA 2.0 may be relicensed under a BY SA 5.0 Chili and a work under BY NC SA 2.0 can be relicensed under a BY NC SA 2.5 Germany.

This allows much better compatibility across versions and jurisdiction licenses. The consequences of the compatibility with jurisdictions versions (3) will be studied in section 3.3. It will be now analyzed how these two changes, Attribution standardization and Share Alike compatibility with later versions, interact and what incompatibilities if any may result from the versioning.

The licenses version 1.0 required derivatives to be licensed only under the terms of this license (1.0) and the licenses version 2.0 and up (2.5, 3.0, etc) accept derivatives to be relicensed under current and later versions, but not under previous versions. Thus, there is no risk that the derivative of a work licensed under a license with the Attribution element could be licensed under a license without the Attribution element.

This change is thus safe in terms in potential source of incompatibility in the situation where only one work is involved, because works under (Non-Attribution) Share Alike licenses may only breed derivatives under similar (Non-Attribution) Share Alike licenses. However, a (Non-Attribution) Share Alike 1.0 work can not be remixed with an Attribution Share Alike 2.0 and up work, because the 4.b. provisions of both licenses are incompatible: 1.0 can be derived and relicensed only under 1.0, and 2.0 cannot be derived an relicensed under a 1.0. Works licensed under a 1.0 license without the Attribution element cannot be remixed with works licensed under any other terms. Thus, the pool of works under a SA 1.0 license is not part of the broader commons which can be reused and remixed with works licensed under more recent versions. To conclude, works under a version 1.0 are not compatible with works licensed under any other version. In that sense, the Share Alike flexibility introduced for version 2.0 was a positive and useful change to avoid this problem in the future, and allow works licensed under different versions to be remixed. But compatibility is limited to licenses carrying the same elements. It has been considered and asked by some users to extend the compatibility to make BY-NC-SA and BY-SA licenses compatible, but the organization has not proceed to that change yet.

c) Link-back attribution requirement

The licensee must attribute the author on each copy, performance or adaptation by conveying the name of the author if supplied, the title of the work if provided, by identifying the use of the work in the derivative, and also, as an upgrade in version 2.0, if it is practically possible to do so, the Uniform Resource Identifier (URI) that the licensor may have provided with the work if it refers to the copyright notice or licensing information of the work.

This additional requirement does not seem to create incompatibilities.

d) Synchronization and music rights

The definition for derivative works is expanded and includes from now on, in case the work is a musical composition or a sound recording, its synchronization with moving images. Music published under a license with the Non Derivative element cannot be mixed with a film because this would be considered a Derivative and not a Collective work. Only music published under BY, BY SA, BY NC, and BY NC SA can be reused to illustrate films and audiovisual works.

This specification creates remix incompatibility in the sense that music under ND cannot be reused to illustrate an audiovisual work. But users who do not read the legal code may be unaware of this legal detail, especially if the music track is used entirely without modification. Synchronization rights are considered as derivative works in US law, but this may not necessarily be the case in all countries. Thus, licensors may be unaware that choosing an ND option will prevent their music to illustrate a documentary. Licensees may be unaware that they cannot reuse ND music to illustrate their documentary, even without modifying the track. An author licensing her music under a BY ND will have her music excluded from the pool of synchronizable music. Besides, if synchronization rights were not considered to create a derivative work in the absence of a more substantial transformation by some jurisdictions, this specification would create incompatibilities between international versions.

e) Limited warranties: the hidden risk of infringement

The most important change between version 1.0 and 2.0 is that warranties are removed from the core of the licenses. Version 1.0 clause 5. entitled Representations, Warranties and Disclaimer would specify that the licensor owns the rights to secure a quiet use by the licensee: the licensor would warrant that the work does not infringe any rights, and that it can be used without paying royalties:

“By offering the Work for public release under this License, Licensor represents and warrants that, to the best of Licensor's knowledge after reasonable inquiry:
- Licensor has secured all rights in the Work necessary to grant the license rights hereunder and to permit the lawful exercise of the rights granted hereunder without You having any obligation to pay any royalties, compulsory license fees, residuals or any other payments;
- The Work does not infringe the copyright, trademark, publicity rights, common law rights or any other right of any third party or constitute defamation, invasion of privacy or other tortious injury to any third party.”

This provision was favorable to the licensee and fostering reuse and remix. Its removal does not directly create incompatibility between works, but at an upper level is a big caveat for the sharing and remix culture legal security. It prevents the peaceful enjoyment of CC works because it may be that CC works may not be used as offered in the license.

In relation to the cascade of responsibility described in the 2.3.3 section, it will be up to infringement procedures and contract law to decide whether a licensor who distributed a work for which she does not own all the rights (either because it contains someone else's work, or because she is a member of a collecting society and cannot offer a work free of charge for all the uses of the grant) can be held responsible if the grant is invalid and the right holder or the collecting society sues the licensee who was expecting to use a "clean" work.

The rationale for the deletion of the warranty presented on CC blog is that warranties can be sold as a commodity. The sustainability of the ecosystem is turned into an optional business model: "licensors could sell warranties to risk-averse, high-exposure licensees interested in the due diligence paper trial, thereby creating nice CC business model."

This issue is also discussed in section 4.2.3 and section 3.3 (as 2.0 France licenses kept the warranty provision of version 1.0 and the Share Alike international compatibility clause will have the effect to remove these warranties after re-licensing of a derivative under a subsequent incremental or different jurisdiction' version). Neither the GNU-GPL nor the GFDL have a clause on representation or the express absence of representation, meaning that authorship is a question of proof left outside the license to be decided through applicable law.

3.2.2 From 2.0 to 2.5 in June 2005[172]

a) Attribution to authors or other parties

Version 2.5 only contains a minor revision: the attribution can be requested to credit the author or any other party: a licensor, a sponsor, a journal, a publisher, an institution. This modification provides more flexibility and freedom to better support more complex and personalized ways and social or scientific norms to request attribution.

For instance, in the case of work for hire, a staff member will be credited for her article, but also the funder, the university and the journal of first publication. It may also help to distinguish the author from the right holder and credit both.

This modification is not expected to create incompatibility, but it increases and expands the protection of the licensor/author attribution rights and creates more burdens for the licensee in order to properly attribute all the necessary parties in the expected way. After version 2.0 standardization of the Attribution element and the possibility to ask to be credited also with a link, it is an additional step toward the recognition of the civil law, romantic version of strong authorship where the author has more strength to exercise her moral right of attribution. Notwithstanding the licenses pending qualification of contractual obligation, in the countries where attribution is weak or does not exist, this may cause the licensees who would not respect the attribution requirement to face a breach of contract, even if the lack of complete and proper attribution would not have been considered a copyright infringement in their

[172] http://creativecommons.org/weblog/entry/5457

jurisdiction.

3.2.3 From 2.5 to 3.0 in February 2007[173]

Versioning to 3.0 has been the biggest revision in the CC licenses history. The process involved the consultation of many partners and stakeholders, including the international affiliates community.

The mention "To the extent this license may be considered to be a contract" has been added in the foreword.

a) Attribution and no endorsement clause

The attribution language has been once more clarified so that a licensor would not imply that the licensor supports or endorses her or the derivative work. This precision is a "No Endorsement" clause, answering a concern by users such as MIT "to ensure that when people translate and locally adapt MIT content under the terms of the BY-NC-SA license, they make it clear that they are doing so under the terms of the license and not as a result of a special relationship between MIT and that person".

This additional specification of the way to express attribution is not creating additional incompatibilities between licenses or works.

b) Compatibility structure between BY-SA and other licenses to be determined

The CC BY-SA 3.0 licenses now include a compatibility structure with licenses to be approved or certified as compatible by CC organization. Once this process hosts other licenses, "licensees of both the BY-SA 3.0 and the certified CC compatible license will be able to relicense derivatives under either license (e.g., under either the BY-SA or the certified CC compatible license)."

This is an extension of the Share Alike interoperability clause. It aims at fostering compatibility through a political decision, rather through an adaptation process such as with the CCi versions of the licenses. It is a progress in the sense that more open content can be mixed with CC BY SA works. However, as license texts are different, it will be studied in section 3.5 what are the possible difficulties raised by compatibility with external licenses, starting with those licenses whose institution started discussion with CC toward compatibility: the Free Art License, and the GNU Free Documentation (GFDL) license. The process, in order to reach full compatibility effects, should be reciprocical: if CC BY SA recognized the FAL as compatible, the FAL should do the same.

[173] http://wiki.creativecommons.org/Version_3 and http://creativecommons.org/weblog/entry/7249

c) Internationalization of the generic/unported licenses

The major innovation of the 3.0 versioning is the internationalization of licenses formally corresponding to US law even if called “generic”. Licenses definitions are now based on international texts and have been renamed “unported”. They are not referring to any specific jurisdiction and are to be ported into the various jurisdictions law of the CCi system. They are drafted with the terminology of the Berne Convention for the Protection of Literary and Artistic Works, the Rome Convention, the WIPO Copyright Treaty, the WIPO Performances and Phonograms Treaty and the Universal Copyright Convention. Rights and subject matter definitions should “be enforced according to the corresponding provisions of the implementation of those treaty provisions in the applicable national law.”

This change brings much clarity and internal coherence to the system, and is not per se creating incompatibilities, which may already exist between jurisdictions’ legislations as it will be further discussed.

d) Moral rights clause, for international harmonization

As the licensor’s right of integrity may be seen as conflicting with the licensee right to make derivatives, CC organization and several jurisdictions felt the need to include moral rights in the wording, though it was already included by some jurisdictions which added this provision during the porting process, or understood it would be applicable by default by the courts because the licenses apply in addition to applicable law. But for more clarification, the provision now appears in both legal code and human readable code. With version 3.0, the unported structure states that moral rights are either retained, or waived and not asserted in jurisdictions where this is possible.

This point will be further discussed in sections 3.3 and 4.2.1, it is likely to create incompatibilities between licenses because the scope of moral rights and the way they can be enforced vary widely from country to country. This incompatibility is not caused by the CC licenses by themselves, but by differences between national laws which are not harmonized.

e) Collecting society clause, for international harmonization

As for moral rights, the language clarifies information which could already have been ported in jurisdictions versions. It describes the situation and law which has been observed by the jurisdictions and confirms that the licensor can waive or not her right to collect royalties under non-waivable and waivable compulsory licensing schemes and voluntary licensing schemes.

This question will be further mentioned in section 3.3 and 4.2.4, it is likely to create incompatibilities between licenses, or to prevent the licenses to work properly, because the scope of compulsory and the way they can be managed vary widely from country to country and affects the ability of licensors to authorize the use of their work for free. This incompatibility not caused by the wording of the CC licenses themselves, but by differences between the two systems. Collective management societies practices are embedded within the

law and within statutory agreements, contracts that rightholders accept to become members of the societies.

f) TPM language clarification

Debian, a prominent organization in the free software community, was concerned by the CC licenses anti-TPM (Technical Protection Measure) clause preventing licensees using the work with technological protection measures which control the access to or the use of the work in a manner inconsistent with the freedom granted in the licenses. The Debian project noticed that the wording would preclude licensees from including CC content on Sony Playstation platforms. They suggested introducing a parallel distribution clause allowing a licensee to distribute the work in any format even protected, provided that the work would also be available in an unprotected format. This possible change has been discussed during the versioning process, but has finally not been included in the 3.0 version because of the opposition of the CCi affiliates community to restrict freedom.

g) Database sui generis rights in CCi versions

Databases were not explicitly included in previous versions of the generic/unported licenses. They are now indirectly covered because the definition of work includes compilation of data to the extend they are protected by copyright law, which vary among jurisdictions. Compilations were already in the definition of Works and thus covered by the licenses, and the difference between a compilation of works and a database of works can be thin.

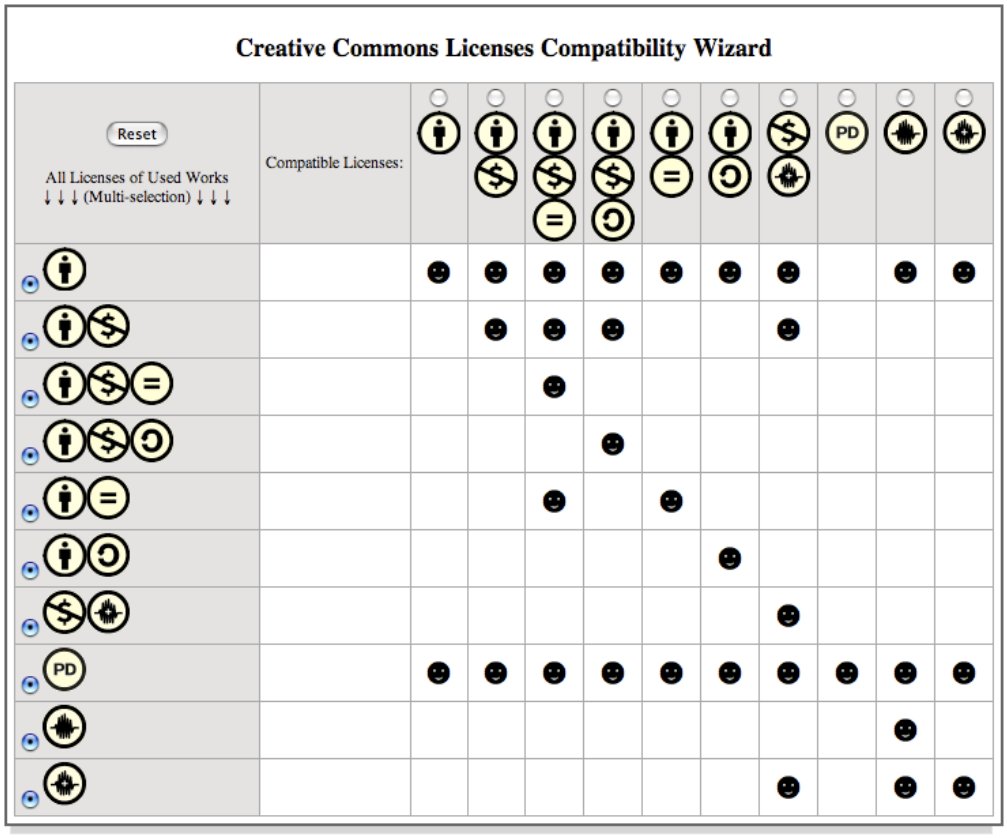
The exclusion of databases is not an actual change within the generic version 3.0, but its mention can be found in the CCi 3.0 porting documentation. Database rights should be waived and the license elements requirements (Attribution, Non-Commercial, No-Derivatives, and Share-Alike) should not be applied to database rights. They had previously been included in several CCi versions (the Netherlands, Germany, Belgium and France among other countries) which added extraction and reuse of substantial parts of a database in the version 2.0 rights grant, as the equivalent to the right of reproduction, performance and distribution for works covered by copyright and neighboring rights.

The goal of this change is to provide clarification to the status of databases in the licenses and interoperability among licenses in different CCi jurisdictions. But it is already a source of incompatibility between licenses because a few of them recognize databases as a subject matter of the licenses and because many databases have been released under a CC license. The topic will be discussed again in section 3.4 and 4.2.2. Databases of works can be distinguished from databases of data or uncopyrightable facts and information, which are now particularly addressed by the CC0 protocol aiming at placing works and other elements as near as possible to the public domain.

3.3 Incompatibility between different options

Offering many options raises information costs and defeats the purpose of the remix culture if they can't be remixed together because of the Share Alike effect.
We will here detail all the concrete impossibilities between options, preventing to remix works under different licenses. SA is incompatible with ND in the sense that no license contains both elements because SA applies to derivatives. Besides that obvious caveat, it is not so easy to list all the incompatibilities, and it should also be noted that the NC clause affects Derivatives but also Collections.

The table below realized by CC Taiwan team helps defining under which license a work and adaptations of it can be relicensed.



Introduction

1. This wizard (chart) above should give you some assistance in figuring out which Creative Commons license you can use to relicense a work.
2. To check out some compatible licenses (i.e., licenses you can use to relicense a work) from licenses of works you are using:
    1. According to those Creative Commons-licensed works you used, check the corresponding Creative Commons license in the left side (vertical-axis) of the chart above.
    2. You can see license names by hovering your mouse (or other point devices) cursor on those deed icons.
    3. Repeat first two steps until all CC-licensed work you used are checked properly.
    4. Alone with your checking process, some smiley faces ☻ may appear in the chart to mark those compatible licenses for each license of works you are used.
    5. For the intersection of compatible licenses, a light-blue background color will appear in the chart above.
    6. You can see the names of intersection of compatible licenses by hovering your mouse (or other point devices) cursor on those deed icons.
    7. This intersection of compatible licenses indicate Creative Commons licenses you can relicense your work under.

8. If there is no light-blue backgrounded cell in the end of your operation, maybe you are using incompatible works.
9. However, you can still look into smiley faces to figure out which work you have to drop out to ensure license compatibility.
10. Then, you can check license compatibility again by using this wizard.
11. Or maybe you can contact the author of particular work to gain extra permissions or rights to use that work.

3. To check out up-stream licenses (i.e., licenses of works you'd like to use in your work) from license you'd like to relicense your work under:

1. According to the Creative Commons license you'd like to relicense your work under, check the corresponding license in the upper side (horizontal-axis) of the chart above.
2. You can see license names by hovering your mouse (or other point devices) cursor on those deed icons.
3. Alone with your checking, some licenses will be highlighted with blue background in the left side (vertical-axis) of the chart above.
4. Those highlighted licenses are usable up-stream licenses comppatible with one you'd like to relicense your work under.
5. You can see those licenses names by hovering your mouse (or other point devices) cursor on those deed icons.

4. By pressing the "Reset" button in the upper-left corner of the chart above, you can clear all selection and re-start again.

Creative Commons Licenses Compatibility Wizard[174]
http://creativecommons.org.tw/licwiz/english.html

The two charts hereafter are part of the CC FAQs section and help to define under which license Derivatives and Collections can be licensed.

Compatibility chart		Terms that can be used for a derivative work or adaptation						
		by	by-nc	by-nc-nd	by-nc-sa	by-nd	by-sa	pd
Status of original work	pd	■	■	■	■	■	■	■
	by	■	■	■	■	■	■	
	by-nc		■	■	■			
	by-nc-nd							
	by-nc-sa				■			
	by-nd							
	by-sa						■	

If I use a Creative Commons-licensed work to create a new work (i.e. a derivative work or adaptation), which Creative Commons license can I use for my new work?

The chart below should give you some assistance in figuring out which Creative Commons license you can use on your new work. Some of our licenses just do not, as practical matter, work together.
The green boxes indicate license compatibility. That is, you may use the license indicated in the top row for your derivative work or adaptation, or for a collective work. The blank rows for the by-nc-nd and by-nd licenses indicate that derivative works or adaptations are not permitted by the license of the original work, therefore you are never allowed to re-license them.

Compatibility chart for derivative works

[174] This application is modified from Licenses Wizard V3.0 of Open Source Software Foundry, and is licensed under the MIT license. The source code of this application can be downloaded here.

Original Work	Commercial Collected Work	NonCommercial Collected Work
pd		
by		
by-nc		
by-nc-nd		
by-nc-sa		
by-nd		
by-sa		

I'm collecting a number of different works together into one resource. Can I include Creative Commons-licensed material?

All the Creative Commons licenses allow the original work to be included in collections such as anthologies, encyclopedias and broadcasts. However, you still have to follow the license the original material is under. For example, material under any of the Creative Commons Noncommercial licenses cannot be included in a collection that is going to be used commercially. The table below will help you work out whether you can include the Creative Commons-licensed material in your collection.
Note that when you include a Creative Commons licensed work in a collection, you must keep the work under the same license. This doesn't mean the whole collection has to be put under the Creative Commons license – just the original work.

Compatibility chart for collections

Creative Commons made the choice to offer several options. This creates internal incompatibilities because not all content licensed under a Creative Commons license is ready to be remixed with other works licensed under another or even the same Creative Commons license.

Open content licenses aim at facilitating the reuse and the remix of copyrighted material by granting clear permissions, and different options are available to suit the needs of a multiplicity of user expectations. What are the transaction and information costs to remix open content material licensed under different, possibly incompatible licenses? What is the impact for users in terms of incentive to reuse works and make derivatives?

This section lists the possibilities between the various combinations and analyzes the unintended and uncertain situations. The diversity of options leads to obvious incompatibilities, unlike some incompatibilities between international versions or with licenses which may be declared compatible which are less visible if not hidden.

Works verbatim reproduction, performance and distribution (without modification)
Works can be copied, performed and distributed only under their license of origin which must accompany each copy or performance.

Collective works
The difference between collective works and derivative works is sometimes unclear and the source of many questions on the various mailing lists.
All CC licenses authorize the inclusion of a work into collective works or collections, to the

extent the work is licensed under the same license, which does not "require the Collection apart from the Work itself to be made subject to the terms of this License". In that case, there is no problem of incompatibility, any CC work may be include in any collection. Even SA works do not require the collection to be licensed under SA terms.
Expectations of virality may be disappointed. But there is one major limitation: works licensed under a BY NC, BY NC SA and BY NC ND cannot be included into a collection which is going to be used for commercial purposes.

Derivative works
BY NC ND and BY ND works cannot be modified. Therefore, they are incompatible with any other work because they cannot lead to derivative works and the question of relicensing of the derivative is avoided.
Only works under a BY license may be remixed with works licensed under any other license and relicensed under any condition, including all rights reserved. BY SA and BY NC SA works can only be remixed and relicensed under the same license.
BY SA and BY NC SA content cannot be combined because of the NC provision: this may be the biggest limitation of the system.
BY NC works can be modified and relicensed under BY NC, BY NC ND and BY NC SA.

According to Katz[175], "incompatibilities between certain Creative Commons licenses may limit the future production and distribution of creative works in ways that today's creators may not intend." He studied effects on the second generation of derivative works made by transforming a first-generation derivative work and how the licenses dynamic may shape the production of derivatives. In his evolutionary model, SA licenses will take more importance, because of their viral effect, but because of the incompatibility between BY SA and BY NC SA, there will be more derivative works released under a BY NC SA license and BY SA works will become isolated and less likely to be reused.

3.4 Incompatibility between different jurisdictions

Creative Commons decided to work with international teams of affiliates. Acting as a network to advise on the project at the international level and to work with national communities, the initial task of the teams is to translate the material and to adapt the licenses to local legislation. For instance, the definitions which have been drafted in reference to international conventions are expected to be replaced by the definitions of national copyright laws. We noted in the previous section that the Share Alike clause admits the relicensing of an Adaptation under a license from another jurisdiction: they are declared compatible. But are they really compatible, meaning that they cover the same subject-matter, offer the same scope of rights and contain the same limitations?

The goal is to foster implementation in order to avoid interpretation problems, and improve compatibility to copyright law. But implementation actually leads to incompatibility with contract law and a consent problem because a Licensor is expected to consent the Adaptation of her Work to be licensed under different, future, unidentified terms.

[175] Katz Zachary, "Pitfalls of Open Licensing: An Analysis of Creative Commons Licensing", *IDEA – The Intellectual Property Law Review*, vol. 46, n°3, 2006, p. 391-413.

We will first present the rational of the Creative Commons porting project before comparing jurisdictions' licenses. We will not analyze and compare systematically all the provisions of all the ported versions of the licenses. On the contrary, we will discuss a few clauses which vary among jurisdictions and which are source of inconsistencies. We selected them either because these clauses raise an issue and/or these jurisdictions illustrate remarkable differences between legal systems. They should allow to assess whether these inconsistencies are a source of incompatibility and jeopardize legal certainty for the first or the second generation of users because of CC choices, or if these differences between licenses which are declared compatible actually do not generate more issues than those raised by the differences already existing in the law because legislations are not harmonized. In other words, is CC creating additional problems to an already difficult situation, or is CC only not solving the cross-national copyright lack of harmonization? For instance, what is allowed under exceptions and limitations to exclusive right and therefore possible even with an ND or an NC license will vary from country to country depending for instance of the scope of their exceptions.

3.4.1 The legal porting

The Creative Commons International (CCi) team coordinates jurisdictions affiliates during the porting process and afterwards, in order to make sure international licenses remain as close to the original version as possible, thus to maintain as much compatibility as possible. International affiliates are expected to provide a re-translation into English of their first draft and share the rational of their proposed legal modifications, which should be kept as minimal as possible.

Over 50 teams around the world translated and adapted the licenses to the language and the legislation of their jurisdictions. With the Share Alike interoperability clause, works licensed under a Share Alike can be remixed with works licensed under a Share Alike license from another jurisdiction, and the resulting derivative work may be relicensed under the Share Alike license of a third jurisdiction. In addition to the compatibility with international versions, the Share Alike clause also foresees a compatibility with a later version of the same license. To add even more complexity, not all the jurisdictions are at the same stage and not all of them translated all the versions. For instance, versions 2.0, 2.5 and 3.0 are available in the Netherlands, while the French jurisdiction is still working with the 2.0 version.

International legal diversity has not been the choice of other free or open licenses systems, which opted for a unique option and jurisdiction, instead of offering a choice in the level of control authors are ready to give in, and local translations, which can not be under the absolute control of a central organization. Creative Commons is the first organization in the open licensing sphere to provide local translations by jurisdiction, which are coordinated but cannot be under the absolute control of the central organization. The latter point is one of the arguments of the Free Software Foundation for not having ported the GNU-GPL and GFDL licenses: while linguistic translations are available for information, they do not have a legal status because the organization cannot be certain of the impact of possible legal differences, notwithstanding errors that may affect localized adaptations.

The purposes of porting the licenses to local laws are numerous. The main advantage of having jurisdiction-specific licenses is to provide linguistic translation for the users, thus

respecting consumer law and fostering acceptability by non-English speaking local communities, and legal adaptation making the interpretation by the local judge easier. Localized texts are more likely to be valid in local jurisdictions that global texts.

Also, teams in charge of the linguistic translation and the legal adaptation are forming a political army of project leads, a form of "political franchising"[176]. These experts are answering to questions by their communities and contributing to the success of the licenses in their country. They are also advising the central organization on best ways to improve the unported licenses and their compatibility with as many legal systems as possible.

3.4.2 Internal validity vs. unexpected inconsistencies

Validity, enforceability or effectivity of the licenses despite possible legal differences from country to country is a goal of the drafters which is expressed in the severity clause: "If any provision of this License is invalid or unenforceable under applicable law, it shall not affect the validity or enforceability of the remainder of the terms of this License, and without further action by the parties to this agreement, such provision shall be reformed to the minimum extent necessary to make such provision valid and enforceable." The effect of this clause is not absolute, in the case of a jurisdiction which contract law would invalidate entire agreements if one clause were invalid.

The desire to ensure internal validity in as many jurisdictions as possible and thus accept differences between national translations justified by differences between national laws in order to be enforceable in the various jurisdictions may have side effects or undesired consequences. Indeed, laws and thus licenses do not have the same definitions for rights and subject matters, and do not address the same concepts. Elements which may be covered by licenses in one country may not be protected in another one, rights may be broader or more limited from one jurisdiction to another jurisdiction. And despite CC and affiliates best efforts to maintain a coherence within the system, a judge may well decide to give yet another interpretation of a concept, such as non-commercial. In that sense, the licenses add complexity to pre-existing multinational licensing issues.

Nevertheless, the Share Alike provision aims at making them all compatible and allows a licensee to license her derivative work under the license of another jurisdiction. The third party C may thus ignore some requirements of the jurisdiction's license chosen by licensor A. There is also a risk that specific provisions chosen by licensor A lose their effects because they will disappear after her work licensed under a SA license will have been derived and relicensed under a SA license from another jurisdiction. Therefore, the validity of the contract is jeopardized because requirements of informed notice may not be fulfilled despite the best efforts to keep the licenses as compatible as possible by minimizing the differences.

These incompatibilities are hidden in the sense that neither licensees not licensors will read

[176] On the structuration on the international community and the relationship between the organization and its international affiliates, see http://governancexborders.wordpress.com/tag/wikimania-preview/ and Dobusch Leonhard, Quack Sigrid, "Epistemic Communities and Social Movements: Transnational Dynamics in the Case of Creative Commons", *MPIFG Discussion Paper* 08/8, 2008. http://www.mpifg.de/pu/mpifg_dp/dp08-8.pdf

the legal code from other jurisdictions. A systematic analysis of differences between clauses should reveal inconsistencies, and therefore potential risks for licensors and licensees expectations and the validity of the agreement because jurisdictions' licenses definition for author, work, rights, restrictions and other conditions will not have exactly the same contractual scope.

This study does not analyze and compare all the 50 versions, but provides some selected examples to demonstrate the contamination risk which may occur from the first generation of derivative works, and grow exponentially after several generations. Examples include the limited warranties and representation, moral rights, the inclusion of related and databases rights in the definition of Work, the scope of applicable rights (what constitutes an Adaptation, what is non-commercial).

3.4.3 Representation of non-infringement

A limited representation by the author is included in several ported versions, but not in the generic 3.0 version. As it has been noted, they were removed between version 1.0 and 2.0, but France 2.0 version kept them for compliance with local law. Thus, any potential French licensee reading the French version, assuming the other jurisdictions' licenses are equivalent and also contain this provision, may expect all CC works to be safe for reuse and free of copyright infringement or other troubles. In the chain of responsibility, it is difficult to know, if the French happen to transmit an infringing work and is sued for that, if she could sue back the original licensor who actually disclaimed any representation. If a work X licensed by A under a US license is transformed by B into a derivative work X, which is re-licensed under the French version of the license, potential licensees C may expect B to carry new obligations that A was not carrying.

Similarly, a contractual limitation of liability arising out of willful or grossly negligent behavior is void according to Section 1229 Italian Civil Code[177]. The disclaimer of liability is thus non applicable in the 2.5 Italian version of the licenses. The New Zealand version has the exact opposite clause: "the Licensor shall not be liable on any legal basis (including without limitation negligence)".

Databases are a subject-matter of the licenses in Dutch, German, French and Belgium versions 2.0 and 2.5. They have been removed from 3.0 (in practice only by the Dutch as the other jurisdictions haven't ported 3.0 yet) and the effects of the optional license elements shall lose their effect and not be applied to databases. Thus, the licensor of a database licensed under a BY SA Netherlands 2.0 license will expect derivatives to carry the Share Alike element and stay in the Commons. However, the Share Alike interoperability clause allows any derivative of the database may be relicensed under a license which may state that the licensing restrictions, including Share Alike, cannot be applied to a database. Therefore, the second derivative will not be shared with the Share Alike element, and the original licensor's expectation will be disappointed as far as BY, NC and SA are concerned: these restrictions will not be applied. It seems difficult to designate a responsible person because the terms of the agreement changed and database rights must be waived according to the Netherlands 3.0

[177] http://mirrors.creativecommons.org/international/it/it-legalchanges.pdf

licenses.

3.4.4 Scope of rights

The scope of applicable rights also varies from one jurisdiction to another. For instance, German law (§31 UrhG)[178] excluded the right to use the work in formats which are currently unknown, and can thus not translate the last sentence of section 3 stating that rights may be exercised in all media and formats whether now known or hereafter devised because such a clause would be invalid in German law and rights would still belong to the Licensor, thus this sentence had been left out the Germany 2.0 version, but later re-introduced in the version 3.0. Italian, Romanian, Greek and probably other copyright laws also forbid transfer of future rights, or rights for unknown types of use of a work. Thus, a licensee reading another version of the license or intending to reuse the derivative version may well think that she is free to transform the work in another new format, without knowing that this prerogative is reserved by the initial licensor.

The non-commercial definition was not translated verbatim by all jurisdictions: for instance, "commercial purpose" may be defined by the Greek judge otherwise than in the unported license. Therefore, if the Greek case law adopts a broader understanding, derivatives of BY NC SA may be used in a manner which may be considered a commercial use in the jurisdiction of the original licensor.

The Canadian version based on Canadian law considers that converting a dramatic work into a non-dramatic work, or adapting it as a cinematographic film, constitutes a mere "Use" and not a "Derivative work". Thus, these usages are authorized even in licenses carrying the ND element. Besides, the moral right of integrity is waivable in Canada and the licenses have included this prerogative in order to ensure that the licensor may permit derivative works. Thus, licensors from jurisdictions with more restrictive moral rights will see the level of protection decrease if the derivative of their original work is relicensed under a Canadian version explicitly waiving moral rights for the subsequent derivatives.

Adaptations are defined quite strictly in Australian copyright law. CC Australia 2.1 ND licenses therefore authorize a number of uses of works which would be considered derivatives in other jurisdictions, for instance the making of a film from a script.[179]

To conclude, the country with the more permissive regime may export risks in more protective or civil law jurisdictions, which nationals may find their expectations disappointed. This adds complexity to international law differences if contracts read by nationals contain different provisions and makes the responsibility even more difficult to locate if the infringer was following the least protective legislation and license.

---

[178] http://mirrors.creativecommons.org/international/de/english-changes.pdf

[179] Bond Catherine, "Simplification and Consistency in Australian Public Rights Licences", (2007) 4:1 *SCRIPTed* 38: "Many of the difficulties in achieving consistency between public rights licences on a global level are a result of the differences in terminology in national copyright laws. (…) The translation of the United States CC licences into Australian law provides a good illustration of the question as to whether national issues must be sacrificed for the sake of international consistency and vice versa.

Differences between licenses jeopardizing contractual certainty are caused by differences between national laws. It seems that CC is taking more responsibilities than it should, and that the ambitious project to endorse the effort of trying to make licenses compatible is a lost cause: CC will not be able to get rid of international differences. Externalizing the interpretation task to the judge is not a brave attitude, but it would have the advantage to not threaten the validity of the Share Alike clause and of the entire contractual chain.

3.5 Incompatibility with other open content licenses

Since version 3.0, the Share Alike clause also declares a compatibility with CC Compatible Licenses. This clause targets open content licenses which are outside the CC system but have equivalent terms and introduces the possibility to re-license derivatives under the terms of other licenses:

"You may Distribute or Publicly Perform an Adaptation only under the terms of: (i) this License; (ii) a later version of this License with the same License Elements as this License; (iii) a Creative Commons jurisdiction license (either this or a later license version) that contains the same License Elements as this License (e.g., Attribution-ShareAlike 3.0 US)); (iv) a Creative Commons Compatible License."

No license has been recognized "Compatible" yet, but discussions have started at least with the organizations curating two licenses, the GNU Free Documentation License (GNU-GFDL) managed by the Free Software Foundation (FSF) in the United States and the Free Art License created by Copyleft Attitude in France. Potentially, all open content licenses[180] could be candidates to that compatibility process.

Efforts have been and are still being led to reach compatibility by inserting a clause in the licenses accepting that derivatives may be licensed not only under the same license but also under licenses which have been recognized compatible. However, discussions are often passionate and results uncertain because communities are ideologically attached to the particularisms of their licensing schemes and not necessarily supportive of the specificities of the other licensing schemes.

As we demonstrated for the compatibility declared between jurisdictions licenses, the Share Alike compatibility is merely a political statement which must be validated by the facts. As different licenses have different phrasing, it should be checked whether differences may also change the content of the grant and its substantial conditions and therefore affect users' expectations and threaten the validity of the consent along the modification chain.

In order to inform the decision of institutions to recognize political compatibility, differences must be scrutinized to see if licenses intend to have an equivalent effect. Besides the uncertainty for licensors, the process requires trust and is all the more controversial that compatibility may be approved also for subsequent versions.

Four methods to improve compatibility between different open licenses and open licensed works will be considered:
- Cross-licensing and reciprocical compatibility *per se* between licenses,
- Combination of works licensed under different licenses and partial compatibility between content,
- Dual-licensing and re-licensing, reaching *de facto* compatibility between content by

[180] IFROSS lists 30 open content licenses at http://www.ifross.org/, click on the "open content" tab and scroll down.

disappearance of one license
- Definition of common freedoms between licenses, one step backwards to go back to the basics.

Each of this method will be presented using the case of one license or ongoing effort to minimize incompatibility between open licenses and works.

- First, the compatibility cross-licensing clause in the Share Alike clause of the licenses, with the example of the Free Art license (3.5.1).
- Second, the provision allowing combination of works licensed under a Digital Peer Publishing License (DPPL, 3.5.2) with content licensed under a CC BY license (which is not compatible because both licenses cover different scope of rights).
- Third, dual-licensing and re-licensing, another option which has been chosen for Wikipedia with the migration from the GNU-GFDL to the CC BY SA 3.0 unported (3.5.3).
- Forth, the definition of a common ground of core freedoms, the standardization path initiated by the Free Culture Definition (3.5.4) to help recognize "free culture licenses".

We will assess the validity and the effect of these different methods to achieve and define compatibility between licenses and works.

3.5.1 Cross-licensing: the example of the Free Art License

Several other open content licenses have terms which are similar to the Creative Commons Attribution Share Alike. However, because of the copyleft provision, works licensed under one license cannot be mixed with works licensed under another close but slightly different license. Even if the intention of the licensors (and to a lesser extend of the drafters) may be similar, works licensed under different open content licenses remain incompatible.

Once external licenses will be recognized compatible, it will be for instance possible to re-license a BY SA work under GFDL, and Free Art License (FAL) works derivatives may be re-licensed under any of the BY SA license CCi versions. Therefore, unintended effects may be demultiplied, as differences between different licenses will add up to differences between jurisdictions' licenses.

Besides the obvious difference due to the fact that similar notions are explained with different words, four main differences exist between the two systems. They will be presented hereafter and their consequence on a potential express compatibility will be analyzed.

First, a practical difference between the CC BY SA unported 3.0 legal code and the Free Art License 1.3[181] (FAL) is that the freedom to distribute the work modified or not is granted provided to the licensee specifies "to the recipient where to access the originals" (article 2.2). This notion is missing in the CC licenses and could be a useful addition in the attribution requirements.

[181] The English translation of the FAL is available at http://artlibre.org/licence/lal/en

Second, the main conceptual difference is the distinction between original copy and subsequent works in the FAL. The inclusion of notion of physical original copy and the concern of its integrity, while authorizing modifications of subsequent works, copies of the original, accommodates plastic arts: paintings, sculptures and installations.
Unlike to the FAL, the CC licenses authorize modifications of the work directly. However, there is no risk that the cross-licensing clause would lead a reader of the CC license modify directly the original of a work licensed under the FAL as the distribution under a compatible license applies to the subsequent work, thus after modifications would have been performed on copies of the original.

The FAL 1.3 clause 2.3 foresees that copies of the original, called subsequent works, can be modified provided that the licensee:

- "indicate(s) that the work has been modified and, if it is possible, what kind of modifications have been made;"
- and "distribute(s) the subsequent work under the same license or any compatible license".

The first sentence requiring describing modifications has its equivalent in the CC licenses and the last sentence, the cross-licensing clause, is comparable to the CC SA compatibility language. However, the recognition of a "compatible license" differs between the two license providers. This is the third substantial difference.

One the one hand, CC as an organization prepared a page to host licenses which will have been recognized "compatible", without however indicating what process or precise criteria should be followed, but by providing a broad, high-level declaration of intent to recognize compatible licenses which have "the same purpose, meaning and effect":

""Creative Commons Compatible License" means a license that is listed at http://creativecommons.org/compatiblelicenses that has been approved by Creative Commons as being essentially equivalent to this License, including, at a minimum, because that license:
(i) contains terms that have the same purpose, meaning and effect as the License Elements of this License; and,
(ii) explicitly permits the relicensing of adaptations of works made available under that license under this License or a Creative Commons jurisdiction license with the same License Elements as this License."

On the other hand, Copyleft Attitude, the organization in charge of the FAL, included compatibility criterias in the text of the license, but without indicating where such licenses will be listed, approved, or, which is an unlikely but possible interpretation, if their inclusion should be deduced by the reader's interpretation of any license regarding the criteria. Those criterias listed under clause 5 "Compatibility" are the following:

"A license is compatible with the Free Art License provided:
it gives the right to copy, distribute, and modify copies of the work including for commercial purposes and without any other restrictions than those required by the respect of the other compatibility criteria;
it ensures proper attribution of the work to its authors and access to previous versions of the work when possible;
it recognizes the Free Art License as compatible (reciprocity);
it requires that changes made to the work be subject to the same license or to a license which also meets these compatibility criteria."

It is rather unclear whether all CC BY SA restrictions under clause 4, but also elsewhere in the core grant, can and will be interpreted as "those required by the respect of the other compatibility criteria". We can assume that both decision processes are still to be refined internally and within the communities.

Will there be a vote from both communities, like Wikimedia Foundation consulted Wikipedians for the Wikimedia migration (see section 3.5.3)? How shall the communities be defined? Unlike to Wikipedians activities which can be registered, thus allowing the foundation to set a minimum limit of 25 edits before a certain date to qualify individuals to be eligible to take part to the vote, there is no registration for individuals or institutions who are using a CC BY SA or a FAL to distribute their works, or who are using CC BY SA or FAL licensed works.
Will there be a public discussion within a defined timeline or until consensus is reached, consensus being defined as the lack of "sustainable technical argument" or "formal objection", like for technical standardization such as ISO or the W3C?

The express compatibility process raises uncertainty and challenges[182], it is very ambitious because it implies to reduce incompatibilities between licenses which have the same objective and therefore reduce the Commons fragmentation. Some decisions to be taken will affect the process:

- The scope of Adaptation (will photos and videogame material be considered Adaptations and not Collections like synchronized music on moving images), and
- The possible extension of the cross-compatibility clause to BY and BY NC SA licenses.

These two last questions have actually been taken into consideration by the drafters of the Digital Peer Publishing Licenses, which are analyzed in the coming section 3.5.2[183].

Finally, the forth difference noted between the CC BY-SA and the FAL addresses related and database rights. Their enforcement should be limited as it should not lead to limit the effects of the rights granted as article 3 states that "Activities giving rise to author's rights and related rights shall not challenge the rights granted by this license. For example, this is the reason why performances must be subject to the same license or a compatible license. Similarly, integrating the work in a database, a compilation or an anthology shall not prevent anyone from using the work under the same conditions as those defined in this license".
If related rights are included in the CC licenses, database rights are waived and not submitted to the BY SA provisions nor to the other restrictions, thus the scope of both licenses vary slightly.

Therefore, these differences should be harmonized before including a cross-compatibility clause.

3.5.2 Combination of works licensed under non-compatible terms: the Digital Peer Publishing Licenses

Digital Peer Publishing Licenses (DPPL) are a set of three licenses (the DPPL, the modular

[182] Jessica Coates, "Playing Well With Others: Increasing Compatibility Between Commons Licences.", Workshop on Asia and Commons in the Information Age, Taipei, January 2008.
http://meeting.creativecommons.org.tw/program:playing-well-with-others
[183] http://www.dipp.nrw.de/lizenzen/dppl/index_html?set_language=en&cl=en

DPPL and the free DPPL) "designed for scholarly content because it covers aspects of authenticity, citation, bibliographic data and metadata, permanent access and open formats"[184].

The basic module of this license, the DPPL, provides rights to use only in a digital format and reserves the rights to distribute the work in printed form. Thus, because of this rights fragmentation, it cannot be considered equivalent and therefore a candidate for compatibility with a CC license which allows to reproduce the work in any format, not only in a digital format. However, it contains a clause entitled Combination with other content:

> **§ 8: Combination with other content**
> (1) The Licensor may combine the Work with other content that may be used under the terms of the Creative Commons license "Attribution" and use the combination, as long as the Work and the other content may still be used separately (e.g. combination of text and photography).
> (2) If the Licensor has combined the Work with other content according to paragraph 1, You may not remove or alter any notice stating that the Creative Commons license applies to the other content and you may not use the Work without the other content. You have to comply with the terms of the Creative Commons license for Your use of the other content.
> (3) You may not use any combination of the Work with other content.

DPPL version 3.0, November 2008[185]

Therefore, a DPPL article may be illustrated by a CC BY photo. This kind of use could have been considered an Altered Version of the Work, any version of the work with changes beyond what the law authorizes. But the combination cannot be further modified or recombined, only one generation of collection is accepted. While it does not facilitate the remix culture, which is not the goal of this open access academic licensing scheme, it will avoid any risk of confusion to decipher further derivatives licensing conditions.

This provision neither does not requires a similar reciprocical clause from CC authorizing CC works to be combined with DPPL works. Indeed, the use of the work in a Collection, the action explicitly authorized by the DPPL, is outside the scope of a CC license. It should be noted that for the purpose of clarification, the text of the DPPL should specify which version of the CC BY license is targeted.

In addition to the rights granted in the first license of the suite (the DPPL), the second license of the suite (the modular Digital Peer Publishing License, m-DPPL) allows authors to decide which parts of their work they will let others modify, these parts will be marked as Alterable Parts (e.g. by a color or highlighting or designation or in the history). Altered Versions should be released under an m-DPPL license and if the modification consists of the addition of a new work, this new work may be licensed under a different license. The §10 provision regarding combination with other content has a final clause stating that if alterable parts cannot be used separately, the entire Altered Version should be released under the m-DPPL while also respecting the CC terms:

> (4) If You combine Alterable Parts of the Work with other content, which may be used under the Creative

[184] http://www.dipp.nrw.de/lizenzen/dppl/. See Euler, Ellen. "Licences for Open Access to Scientific Publications - A German Perspective." *INDICARE Monitor* 2, no. 4 (2005). http://www.indicare.org/tiki-read_article.php?articleId=117

[185] http://www.dipp.nrw.de/lizenzen/dppl/dppl/DPPL_v3_en_11-2008.html

Commons Licence "Attribution", in such a way that the Work and the other content cannot be used separately (e.g. insertion of text into other text), You are obliged to grant the right of Use for the entire altered version of the Work under this Modular DDPL Licence to anyone exempt from charges and in addition You have to comply with the terms of the Creative Commons Licence.

m-DPPL License Version 3.0, November 2008[186]

Again, this provision does not require a reciprocical clause from CC, as Collections don't need to be CC licensed. Collections are not submitted to the Share Alike effect which "applies to the Adaptation as incorporated in a Collection, but this does not require the Collection apart from the Adaptation itself to be made subject to the terms of the Applicable License.[187]

However, there might be some difficulties if changes towards an Altered Version happen to lead to an Adaptation rather than a Collection, as it might be the case if both parts cannot be used separately as defined in §10 clause (3) of the m-DPPL. In that case, the Share Alike CC provision would require the Adaptation to be released under a "compatible" license, while the m-DPPL would require the Altered Version to be distributed under the m-DPPL. This scenario is clearly an unsolvable incompatibility.

Also, §8 (3) of the m-DPPL states the work can be combined with content provided under the CC license or the GNU GFDL (again, versions are unspecified) under the conditions mentioned above, but the GNU GFDL is not further mentioned in §10.

The third license of the project, the Free Digital Peer Publishing License (f-DPPL)[188], is closer to the copyleft spirit than the two other licenses, as it allows publishing the document not only in digital format but in any media, and requires to distribute the modified document under the same conditions. Thus, despite some additional provisions regarding integrity and citation, it is closer to being a potential Compatible License with the CC BY SA 3.0 than the DPPL and the m-DPPL.

The f-DPPL §10 provision regarding combination with other content is similar to the aforementioned clauses of the DPPL and the m-DPPL, and contains a fifth final clause stating that if the work is combined with a work licensed under the CC BY SA license or the GNU GFDL, the new work (e.g. the collection in CC terminology) should be licensed under a CC BY SA or GNU GFDL (versions are still missing):

(5) If You combine the Work with other content, which is provided under the Creative Commons Licence "Share Alike" or the GNU Free Documentation License, for combined Use, the new Work may only be Used under the terms of the Creative Commons Licence or the GNU Free Documentation License.

f-DPPL License Version 3.0, November 2008[189]

This unilateral compatibility clause makes it possible to have Collections of DPPL and CC works. It is not necessary to incorporate such a clause in the CC licenses, as the Share Alike clause does not apply to the collection incorporating the Work besides the Work itself.

[186] http://www.dipp.nrw.de/lizenzen/dppl/mdppl/m-DPPL_v3_en_11-2008.html
[187] Article 4b of the CC BY SA 3.0 unported license, http://creativecommons.org/licenses/by-sa/3.0/legalcode
[188] http://www.dipp.nrw.de/lizenzen/dppl/fdppl/f-DPPL_v3_en_11-2008.html
[189] http://www.dipp.nrw.de/lizenzen/dppl/fdppl/f-DPPL_v3_en_11-2008.html

However, there is no such compatibility clause for Altered Versions, the equivalent of Adaptations in CC terminology. The f-DPPL only avoids incompatibility with CC BY-SA (and GNU GFDL) works of (f-DPPL licensed) works incorporated in Collections and resulting Collections, but does not handle works as modified in Adaptations (to the extent that collections and adaptations in CC terminology are equivalent to combinations and altered versions in DPPL definitions, which is uncertain).

3.5.3 Dual licensing and re-licensing: Wikipedia and the GNU-GFDL

Dual licensing designates the action to license one's work under two different licenses. Multi-licensing involves more than one and potentially more than two licenses. For the sake of simplicity, we will address dual-licensing only. As explained in the CC FAQs, dual licensing does not mean that the provisions of both licenses will apply simultaneously, but that the licensor gives the choice to the public to apply one or the other. The purpose is twofold, tend to avoid or minimize license incompatibility issues by providing users more choice to reuse and incorporate one's work, and segment market categories to allow multiple business models, for instance by giving more rights to non-commercial users, initially as this practice comes from the software industry, offer for free under the GNU-GPL or for a fee under conditions which are compatible with proprietary software.

However, the risk of dual-licensing is to postpone compatibility issues and add further complexity. Indeed, a user may eventually stop dual-licensing and choose one or the other license to distribute her derivative, which will thus be no longer compatible with its original work: it is impossible to merge back children into parents. Besides, it introduces complexity, as it may be difficult to assess what part of a composite work will be under what license, for instance heavily edited Wikipedia articles.[190]

Nevertheless, an ad-hoc dual licensing solution has been defined to accompany the migration of the Wikipedia project from one licensing scheme (the GNU-GFDL) to the CC BY-SA 3.0 unported. The objectives to move to a CC license are twofold:

- To avoid some of the inconvenient requirements of the GFDL, mostly in terms of attribution and notice requirements, and
- To allow compatibility with other large projects using CC.

The collaborative encyclopedy Wikipedia started to use the GNU Free Documentation License (GFDL) and wanted to switch to the CC BY SA which was not available at the time the project started. The method which has been applied differs substantially from the SA cross-licensing clause. The GFDL actually allowed projects to change their licensing terms and Wikipedians voted in favor of the change. The procedure involved is obviously even more questionable than the previous issues we considered so far regarding the consent of licensors[191]. It adds incompatibility issues as it also opens up to incompatibilities with

[190] See "Why not dual license" at http://meta.wikimedia.org/wiki/Guide_to_the_dual-license and "The case against multiple licenses" at http://en.wikipedia.org/wiki/Wikipedia:Multi-licensing, last visited 05-02-2010.

[191] Shaffer van Houweling Molly, "The New Servitudes", *Georgetown Law Journal*, Vol. 96, p. 885, 2008. Wikitravel also seeked community consensus for more compatibility with other projects using a CC 3.0 license, but needed to upgrade from CC BY SA 1.0 which did not even have a compatibility mechanism with subsequent versions in the SA clause.

jurisdictions' versions.

The GFDL has originally been drafted for software documentation. Its requirements in terms of attribution and invariant sections are very demanding and that difference between the GFDL and the CC BY SA licenses making it easier to attribute in the CC system, as well as the desire to foster compatibility with other projects using a BY SA, justified the need to change and the choice of CC. The migration process led to numerous discussions to ensure the consensus of the community, if not the consent regarding contract law validity requirements, including the definition of free cultural works and a statement of intent by CC[192].

3.5.4 Free Culture core freedoms: defining Open License

Instead of considering all the legal and policy differences between licenses making it difficult to cross-license, dual-license or re-license works, their derivatives and collections thereof, thus weakening the commons, another intellectual path is to compare licenses to extract common points, or most relevant clauses, in order to define the substance of an Open License by a series of shared principles.

The work led by the FSF[193] and the OSI[194] to define Free, Libre and Open Source Software as well as the definitions of Free Cultural Works[195] and Open Knowledge[196] are a source of inspiration toward the definition of such principles.

Defining core freedoms or principles helps reaching consensus between communities of licenses which aim at becoming compatible though a cross-licensing clause: such a process helps to compare the licenses and to commit to stick to the principles when versioning after promising compatibility would be maintained.
On a more theoretical level, it allows understanding what is exactly at stake and what are the needs dictated by copyright and usages limits to open up a work.
On a practical level finally, it could help to reduce the number of options and the complexity of licenses wording. It could even constitute a human-readable version, or a short readable license.

Several core notions are to be studied across the various licenses and definitions available: the level of attribution and notice requirements, as well as the admissible but non necessary restrictions (such as the Share Alike effect) and the non-admissible restrictions which should be excluded, for instance reserving or preventing specific usage purposes (commercial use,

---

[192] Approved for free culture works: http://creativecommons.org/weblog/entry/8051
Statement of intent for BY-SA: http://creativecommons.org/weblog/entry/8213
[193] The Free Software Definition contains "four essential freedoms" and provides interpretations of what they include and do not include: http://www.gnu.org/philosophy/free-sw.html (see also Why Open Source misses the point of Free Software at http://www.gnu.org/philosophy/open-source-misses-the-point.html
[194] The Open Source Definition criteria are here: http://www.opensource.org/docs/osd and a commented version provides the rational here: http://www.opensource.org/docs/definition.php
[195] The definition of Free Cultural Works is available at http://freedomdefined.org/Definition
[196] The Open Knowledge Definition, addressing not only works but also data and government information, is at http://opendefinition.org/1.0/.

derivative works, technical restrictions[197]…).

Freedoms: Rights to Use
An open license grants all the necessary rights to access, copy, perform, distribute and modify a work, including in a database, a collection or a modified version and all types of usage.
The work and its source should be legally and practically accessible and modifiable.

Admissible conditions: Credits, Notice and Metadata
The author may require the work to be accompanied in an unmodified way by:
- the name, URL or a link to the text of the license,
- the title of the work, attribution information (author, performer, other right holder, sponsor…) as well as modification history of the work to the extend they are provided in a reasonable way according to standards of citation,
- digital signature, original source or location and other metadata.

Non acceptable restrictions: Legal, Technical and Economic Usage Restrictions
An open license should not accept or impose:
- legal restrictions on the exercise rights to limit the users who may exercise the freedoms or the territory, scope, domain or field of usage,
- technical restrictions to access, download and edit a digital copy of work (technical protection measure, compulsory registration, distribution in a non-copiable or non-editable format…),
- economic restrictions to access and copy a digital copy of the work (distribution for a fee, in a format which is not free of charge…)

Open licenses core freedoms and restrictions: a synthesis

This subjective synthesis of the provisions composing an Open License tries to provide a standard of freedom and to suggest the amount of rights and conditions which are necessary to open up a work. It may help to compare the licenses among each other in the process of reaching compatibility among each other. Shorter than the Free Culture and Open Knowledge Definitions and building upon them, it can also be the starting point of a Social Contract or Guidelines *à la* Debian[198], a "set of commitment"[199] at the basis of a definition for open licensing.

[197] On technical restrictions, see Dulong de Rosnay Melanie, "From Free Culture to Open Data: Technical Requirements for Access and Authorship", in Danièle Bourcier, Pompeu Casanovas, Mélanie Dulong de Rosnay, Catharina Maracke (eds.), *Intelligent Multimedia. Sharing Creative Works in a Digital World*, Series in Legal Information and Communication Technologies, Vol. 8, European Press Academic Publishing, Florence, 2010, p. 47-68.
[198] Debian Social Contract, version 1.1, April 26, 2004. http://www.debian.org/social_contract
[199] *Ibidem*.

4. Impact of the differences between licenses

The validity of the contract may be jeopardized by two elements affecting the consent of the parties: who and what? The definition of the parties (section 4.1) and the scope of rights (section 4.2) are indeed essential pieces of information to build an agreement to allow informed consent, an important condition in contract law and to authorize the making of derivative works which would have constituted an infringement without the license, an important feature of open licensing. What rights and subject-matters are exactly covered? Is the legal code clear? Are all the legal codes clear and licensing the same rights and subject-matters? Isn't the human readable deed misleading on that point?

Besides the fact that other rights than copyright, such as publicity rights or privacy are explicitly not covered, do people know if the license covers the entire subject-matter which may be subjected not only to copyright defined strictly, but also to neighboring and sui generis rights, or is the scope of what is covered by the license the first uncertainty?

After detailing the external and internal incompatibilities and inconsistencies, we will now evaluate their actual impact on contract formation and on the ability to make derivative works. Some consequences may be theoretical, minor or harmless, while some others may seriously endanger the validity and the enforceability of the system in some jurisdictions at least, including the ability to make derivative works. Before considering possible solutions to improve the system, it matters to assess whether correctives are really necessary. Indeed, if there is a severe incompatibility and substantial cases where the licenses cannot be held valid and enforced, it could be dangerous using them, at least not worth it. The impact could be that licensors may not be able to require their conditions to be enforced, and that licensees may not be able to claim the benefit from a grant which is more generous than copyright law, thus spreading probably involuntary infringement and creating obstacles to the mash-up culture.

This section will focus on the hidden risks of external and internal inconsistencies, rather than on the visible incompatibilities, and assess actual consequences for users of the system.
Thus, we will not further analyze the impact of the differences between options. On the contrary, we will focus on the consequences of the differences and incompatibilities which may jeopardize the validity and the enforceability of the agreement.

The differences between the licenses may cause confusion, they may also endanger the validity of the agreement if the rights granted are not the same for all parties, and lead to involuntary copyright infringement.

It will be assessed what rights are at the entrance of the licensing process (when a Licensor licenses a Work) and at the exit (when a Licensee obtains that Work and wants to redistribute it or to make a derivative and become a Licensor). A logical principle is that it is not possible to license more rights than one owns. Licensors cannot license more rights than they own, and licensees cannot enjoy (and then further distribute or license) more rights than they were actually granted. Thus, if rights are not the same for all the parties because of differences hidden in the licenses different versions, there is a problem.

First, because parties do not agree on the same subject-matter, the agreement itself may be invalid if the contract cannot be formed because the object is not clear.
Second, if a condition is deemed stated by one party but hidden to the other party, this will cause unvolontary infringement and endanger the ability to share and remix. The impact will be demultiplied along the chain of derivatives as the Share Alike clause allows to use yet another license recognized compatible but in reality different.
This section considers practical and theoretical issues related to the ability to use and modify works licensed under conditions which present differences. Not only licensors and licensees who create, distribute under Share Alike terms and modify works can be affected, but also service providers and intermediaries licensees which simply broadcast or synchronize a musical work.

4.1 Identification of the parties and enforcement

Who are the parties? Are they clearly defined by the legal code? Are they identified? Do they exist? Are they capable parties?

Does the license give the possibility to identify the rights owner? Following the analysis of the licenses main clauses led in section 2.2.3, according to which law or international convention are the rights defined? The unported text is not directly enforceable because it uses the vocabulary of international conventions which take effect and are implemented in jurisdictions; still, the unported version is more used than jurisdictions versions ported by the international project leads because it is available earlier and maybe also because it gives an impression of worldwide enforceability for international projects. Is this unported text thus really relevant and appropriate for public use, or should it be reserved for internal porting purposes, maybe as a matrix for international projects leads?

Are they legally entitled to license the work? We already noted (in 2.2.3.) that the Licensor is not identified and that there is a confusion between Rightholder and Licensor.
Therefore, enforcement may be difficult if parties are unknown and no further information is available on the website or with the attribution elements. Similarly, enforcement is threatened if the licensor is not an authorized party (or if she does not own sufficient rights, see further in section 4.3.4 on the absence of representations of non-infringement).

Subsidiary, it can be that a Licensor is a minor. Can the parties to a copyright-related agreement and a contract be minors? Is there a need for parents authorization? If not, is the license still binding? In principle, minors are incapable and the contract would be void, but children enter into standard agreement all the time for instance when buying a train ticket.

Licensors are committing for the entire duration of copyright, thus after their death. Can they commit their heirs? Can heirs change the licensor's mind and revoke the license, thus affecting licensees? The question of inheritance should not threaten the balance of the system, as the heir inherit if the author had not previously disposed of the rights in favor of a third party[200], but the existence of such a principle should be checked in other legal systems.

[200] Principle §4 (2) (b) in Study Group on a European Civil Code/Research Group on EC Private Law, von Bar Christian, Clive Eric (ed), *Principles, Definitions and Model Rules of European Private Law - Draft Common Frame of Reference (DCFR)*, op cit.

Similarly, bankruptcy opens up to the possibility of revocation if the licensor was a company and its assets are sold.

The enforceability of the license is a crucial point. If the licenses are not valid, they cannot be enforced. Even if they were valid, but not enforceable, they would be legally worthless, as neither licensors nor licensees could seek injunctions and/or remedies in case provisions are not applied by another party. Licensors would not be able to require their works to be reused under the same conditions, and licensees would be unable to benefit from a grant which is more generous than copyright law.

Not identifying the licensor and the rightholder does not help to start an action if the infringer is unknown or incapable. Nevertheless, until now, all case law examples demonstrated that the licenses were held enforceable by both licensors[201] and licensees[202] and in both civil and common law jurisdictions, which is a good sign. Additional case law may help better determine who can claim what, based on what ground and which applicable law. However, case law only is not a sign of enforceability, as many cases of infringement never go to Court precisely because parties cannot be identified. An example of clause which is frequently violated is the Non Commercial restrictions, and licensors can practically not contact all the blogs which reuse their works with commercial banners because they are not reachable parties.

4.2 Scope of rights granted

The differences between the scope of right have consequences on the formation of the contract if there is no agreement of the object, and on the ability to make derivatives if hidden differences may also hide that an action will constitute an infringement in one of the licenses versions but not the other.

The differences between the licenses scope of rights may be due to the fact that the Commons Deed does not include all the rights mentioned in the Legal Code, for example the difference between an Adaptation and a Collection (4.2.1). They may also be hidden in the jurisdictions' versions (4.2.2), which is more dangerous because jurisdictions Legal Codes are declared equally binding and valid. Differences in the scope of rights actually granted (according to Licensors and Licensees who consented to different jurisdictions licenses) which will be analyzed are the following clauses or absence thereof in the licenses: database rights (a.), moral rights (b.), representations of non-infringement (c.) and collecting societies (d.).

4.2.1 A difference between formats: Collections and Adaptations

It was noted in section 2.2.1 that the notion of work should be properly defined in the notice sentence in order to know to what item the license applies. One notion which is not reflected in the Commons Deed but has consequences explained only in the Legal Deed is the

[201] e.g. Case Jacobsen in the US, op cit.
[202] e.g. Case EDU 4 in France, op cit.

difference between a Collection and an Adaptation. The difference between a Collection and an Adaptation is a legal matter which is not transparent to the laymen. Still, the Share Alike clause does apply to Adaptations, but not to Collections:

"This Section 4(b) applies to the Adaptation as incorporated in a Collection, but this does not require the Collection apart from the Adaptation itself to be made subject to the terms of the Applicable License."

However, the Commons Deed sentence could imply that it apply to both transformative items because it does not define and target to exclude Collections as obviously as legal deeds do:

"If you alter, transform, or build upon this work, you may distribute the resulting work only under the same or similar license to this one."

Therefore, a Licensor could expect the Licensee reusing her work in a Collection to be bound by the Share Alike clause. Similarly, a Licensor could expect the synchronization of his music on moving images to be a Collection if the song is unmodified, used in its entirety without any cut. However, the Legal Deed explicitly considers this use as an Adaptation. Therefore, a Licensor might in good faith reuse a music track under a ND license, depending of her understanding of the action of "building upon".

The question of the lack of certain elements in the human-readable Commons Deed can have two interpretations: either it hides some information and may invalidate the consent, or it is not a binding document anyhow and only the legal deed will be interpretated and applied[203]. In principle, only the Legal Deed is binding, but how binding can it be in practice if people read only the Commons Deed? This problem affects in general browse and click-wrap and standard form contracts that nobody reads. The Commons Deed even if it does not contain all the information of the Legal Deed, will provide at least some information. However, differences hidden in jurisdiction's version have a greater impact on the informed consent and thus the validity of the agreement.

[203] Garlick, Mia. "Creative Humbug? Bah the Humbug, Let's Get Creative!" *INDICARE Monitor* 2, no. 5 (2005). http://www.indicare.org/tiki-read_article.php?articleId=124: "Much of what is in the Legal Code is not in the Commons Deed (or the metadata) and no doubt, all legally untrained people who use the Creative Commons licenses and/or works licensed under a Creative Commons license are thankful for this. For example, neither the "Warranties, Representations & Disclaimer" clause, nor the "Limitation on Liability" clause, nor the "Severability" clause nor the "No Waiver" clause are included in the Commons Deed or the metadata. These clauses – whilst necessary to construct a legal document – do & arguably should (for the sanity of the general public) remain the preserve of lawyers and the courts to argue about and interpret."

4.2.2 Differences between jurisdictions

How informed can be one's consent to have her work being later adapted and licensed under a different jurisdiction's license by virtue of the Share Alike 2.0 and 3.0 compatibility clause? And what about the reciprocical action, if one want to adapt a work which has been licensed under a license of the Japanese jurisdiction, how can one understand to what one commit? Even if re-translations into English and an English explanation of substantive legal changes are made available on a section of the CC website[204], all jurisdictions licenses will never be accessible in the language of the prospective licensor and licensee. Variations contained in future versions (3.2), in jurisdictions versions (3.4) and in future versions of future compatible licenses (3.5.1) cause legal insecurity. How can one be bound by something one didn't have the opportunity to agree to?

Parties consent to one legal code, but cannot consent to all the other legal codes under which their modified work may be relicensed after the Share Alike compatibility clause, because they are not accessible pieces of information. The proliferation of licenses and related information costs are jeopardizing informed consent. Too many licenses and an increasing complexity make it impossible to be notified of and understand all the possible future terms of agreement for both licensors and licensees. If there is no meeting of minds, no agreement will be validly formed, and it would be pointless to attach a license to a work if it is not a valid contract.

This caveat on the validity of the Share Alike compatibility clause endangers the sustainability of the system. The initiative to have localized versions of the licensed to foster their enforceability may actually be counterproductive. And the cross-licensing and re-licensing efforts may be useless also if they invalid the agreement because Licensors could not consent the derivative of their work to be relicensed under conditions they were not aware of, and even if the agreement would be held valid, Licensees may infringe Licensors rights as the scope of rights granted is not the same.

Rights which differ between versions and which will be analyzed in this section are the following: database rights, moral rights, absence of representations of non-infringement, provision on collective societies. These four examples illustrate differences between jurisdictions, but also between subsequent incremental versions.

a. Database rights

The scope of rights may vary from jurisdiction to jurisdiction, as noted in section 2.2.3 analyzing the clauses. One specific right even varies among versions of the licenses. Databases are a subject-matter of sui generis rights in European jurisdictions, which grant

[204] By clicking on each of the flags of the projects at http://creativecommons.org/international/, for instance http://creativecommons.org/international/ar/ for Argentina leads to a comparison between the unported and the ported versions based on the local legislation: http://mirrors.creativecommons.org/international/ar/english-changes.pdf

specific rights on the database, in addition to rights on the copyrightable elements which constitute the database and the database itself if selection and arrangement are original[205].

Sui generis database rights have been integrated in the initial porting by several European jurisdictions in 2004[206], because they are part of the applicable legal framework surrounding the use of copyrighted works, subject-matters of the licenses. Copyrighted works can be gathered within databases of works, and it has been considered as useful by several international projects to allow rightholders to distribute databases with more freedom, and allow the public to "extract and reuse" beyond legal exceptions and limitations.

However, they have been explicitly taken out of the scope of rights licensed by version 3.0 as mentioned in section 3.2.3, in order to fulfill the needs of the scientific community regarding databases of data. Science Commons, the initiative of Creative Commons dedicated to science, demonstrated that applying license elements (BY, NC, ND, SA) to scientific databases is not recommended for science because the flow of information should be unrestricted and also because it is difficult even for specialized lawyers to distinguish what part is a database and assess what is a commercial use[207]. In order to avoid some complexity, the database sui generis right is part of the subject-matter (the definition of Work include databases) and of the license grant, but it is waived and not subjected to the restrictions included in clause 4 before the collecting societies and moral rights language. But as a side-effect, database right is not submitting to the clause preventing to distribute the work with a technical protection measure. Thus, it is unclear whether the waiver of the database rights by the producer and the restriction to apply a TPM on the individual works would also prevent the use of a TPM on the database. Could works licensed under a CC 3.0 license, but contained in a database which is not licensed under a CC license, be impossible to download conveniently as a whole because even if the right to extract substantially has been waived, the use of a TPM is not excluded?

Also, the exclusion of database rights makes it impossible to reserve commercial rights on the use of a database, which can disappoint both potential licensees and licensors. Isn't there a need to maintain databases in the scope of the licenses because they address not only databases of data, but also databases of copyrighted works, and because there is such an applicable right in some jurisdictions? Is the CC0 protocol fulfilling such a need?

An argument to exclude database rights from the scope of the CC licenses was the risk of exporting such a protection into jurisdictions which do not recognize a legal protection to databases, such as the United States. Does the Share Alike international compatibility clause have such an effect? If yes, does it really disappear after version 3.0 or is it too late to fix the problem? Can database owners still use a 2.0 license from the French jurisdiction if it is the

---

[205] Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases, ECOJ L 077 , 27/03/1996, p. 20-28.

[206] It is also part of the grant of the f-DPPL based on German law: "This license agreement shall further entitle You to incorporate the Work in electronic databases or other collections. Should You attain Your own rights to databases or collective works, You may not use these to restrict or prevent further Use of the Work.", f-DPPL clause 2 §2 (2), http://www.dipp.nrw.de/lizenzen/dppl/fdppl/f-DPPL_v3_en_11-2008.html

[207] Comments on the Open Database License Proposed by Open Data Commons by Thinh Nguyen, Science Commons Reading Room. http://sciencecommons.org/resources/readingroom/comments-on-odbl; Protocol for Implementing Open Access Data, http://sciencecommons.org/projects/publishing/open-access-data-protocol/; FAQ about the Database Protocol, http://sciencecommons.org/resources/faq/database-protocol/

only version available in that jurisdiction before the release of version 3.0 and even after? If from version 3.0, databases are not subjected to CC conditions, what is the status of databases which have already been licensed? Finally, what is the status of subject-matters which have already been licensed under a CC 3.0 license and happen to be databases, even if their licensors were not aware of the distinction between legal categories and were expecting the restrictions to apply to their creation as a whole, is the license invalid because the intended subject-matter does not match with the targeted subject-matter?

It seems that the removal of database sui generis rights was a decision which went beyond fulfilling its initial goal and which left many questions unanswered, in particular the impact on databases of works in jurisdictions where such a right exist and where licensors might want to waive it in order to fully open up their creation.

b. Moral rights

The moral rights considered in this section are essentially the right of attribution and the right of integrity, to the extend they may create incompatibilities by affecting the making of derivative works. In addition to threatening the making of adaptations and creating involuntary infringement, the question also targets the issue of consent because moral rights standards vary and some of them have been embedded inside the license, sometimes to explicitly waive them as it is the case with the 2.0 Canada licenses waiving the right of integrity, but possibly to incorporate into the agreement. Indeed, moral rights are deemed not affected by the license from the Commons Deed level. Technically, it means that the space of freedom to make derivatives even from ND-licensed works will be broader in jurisdictions which have weaker moral rights than in jurisdictions which have stronger moral rights.

An example of international differences considering moral rights should be chosen in order to illustrate that jurisdictions versions may have different expectations, jeopardizing the validity of the agreement but also the ability to make derivative works if one derivative is considered an infringement of moral rights in one jurisdiction but not in the other.

In Common law countries and especially in the United States, moral rights are often considered as a threat from the civil law tradition jeopardizing the normal exploitation of works, fair use and the remix culture. However, it can also be argued that the CC licenses express the will of the author and are an embodiment of her rights to control the use of her work by dedicating it to the commons[208].

French law is a demanding standard regarding moral rights and can be used an example to illustrate possible problems which may arise from the Share Alike compatibility between jurisdictions versions which are different.

French law[209] grants four categories of moral rights to an author who may neither license, transfer or abandon these rights as they are “perpetual, inalienable and imprescriptible”: the right of paternity or attribution, the right to the integrity and the respect of the work, the right

[208] CC France FAQ on the compatibility of the licenses with French moral rights provisions: http://fr.creativecommons.org/FAQjuridiques.htm
[209] Article L.121 of the French Intellectual Property Code.

of disclosure and the right of withdrawal. In a nutshell, CC non-revocability provision triggers the right of withdrawal[210] and the right of disclosure, the right of integrity, or the right of respect is questioned by the CC licenses authorizing modifications in advance without having reviewed them, while CC attribution provisions can be interpretated as fulfilling the moral right of attribution, which we will start with.

The moral right of attribution seems fulfilled by CC provisions which require specific crediting of the author, indicating the title of the work and the modifications[211]. Authors are expected to properly indicate on their work or on their website the license and their name or additional information as they wish to be credited. They should also be specific about what is being licensed: only the text or the images of a website, also the graphics, the lyrics but not the music of a song, etc… Indeed, when users will redistribute or adapt their work, they should be able to understand what is being licensed and to fulfill the requirement requested by the licenses:

- Continue to indicate the license when distributing or performing the work, in order to inform others of the conditions under which the work has been made available by its original author,
- Attribute the original author in the way she wishes and explain that, for instance, the new work is a translation.

Incorrect attribution is jeopardizing the reusability of works and the making of derivatives, the consent of the licensors and the legal certainty of the licensees. It can lead to both breach of contract and copyright infringement. Licensors and licensees should follow best practices[212] for marking and crediting works in different formats.

The enforcement of the moral right of integrity seems less problematic as distortion, misrepresentation and modification of context are in theory handled by the attribution clause specifying that modifications must be identified:

"Adaptation, including any translation in any medium, takes reasonable steps to clearly label, demarcate or otherwise identify that changes were made to the original Work."
in the case of an Adaptation, a credit identifying the use of the Work in the Adaptation
upon notice from any Licensor You must, to the extent practicable, remove from the Collection/Adaptation any credit
You may only use the credit required by this Section for the purpose of attribution in the manner set out above and, by exercising Your rights under this License, You may not implicitly or explicitly assert or imply any connection with, sponsorship or endorsement"

However, attribution information is often incomplete, or does not properly follow the work and its subsequent derivatives. Also, when credit is removed on the demand of the original author, how can such information be displayed again at a later stage if the author wishes to be attributed again? This scenario is not a legal fiction, but a requirement for those countries where attribution cannot be perpetually abandoned, and for the cases where the derivative of the derivative makes honor to the reputation of the author, while she did not appreciate the first derivative and did not wish to be associated.

---

[210] Which requires the indemnification of the other contracting party and is (almost) not exercized, therefore the risk is more theoretical.
[211] And of what requirements are reasonable in order to avoid a misuse of moral rights by overreaching clauses, also in relation to the requirement of license notice to be conveyed with each copy of the work, the credit removal clause and anonymity in case an author wants to be credited again after derivatives have been created and distributed without crediting her.
[212] http://wiki.creativecommons.org/Marking

The fact that the rights granted must be exercised with respect to the moral right of respect of the author (or performer) who may oppose distortion or mutilation that could be prejudicial to her reputation cannot be further regulated by the licenses, it is a matter of national legislation enforced by the judge. An author could exercise her moral right against a certain use of her work, its reproduction in a particular context, or a modification and then seek injunction or damages against third parties who would have incorporated the incriminated work. But it should be noted that this right is not absolute. Court have a margin of appreciation to weigh the interests at hand, which limit the risks to see moral rights applied for patrimonial reasons by one party limiting the freedom of expression of the other party. Besides, a judge could argue that claiming moral rights after authorizing modifications is bad faith, and disregard the complaint as abusive. Finally, damages for such cases are often symbolic, another argument to demystify the risk of moral right of integrity. But injunctions preventing the further distribution and commercialization of the work are rather common and the impact of the moral right of respect is indeed to jeopardize the use and reuse of CC works in those jurisdictions where it may be applied.

c. Representation of non-infringement

Authors have to consider various questions before deciding to apply a Creative Commons license. The licenses are based on copyright, and are thus applicable on copyrightable works only. According to the FAQs, despite the absence of warranties, potential licensors have to make sure that they own the rights they are about to license to others, otherwise they might transmit a junk work which will jeopardize the legal certainty of those who will reuse it. Potential licensors may have to ask the permission of possible co-authors, authors of pre-existing works, employers, or previous assignees such as collecting societies before applying a CC license. Besides, not all the rights that might be contained in a work are licensed in the grant: for instance privacy or publicity rights of the subjects represented in a photography may object to the use of their image. The CC license will cover the copyright of the photographer, but a separate agreement should be negotiated to cover publicity rights.

Two points which may invalidate, or at least reduce the interest and the value of the license grant in its substantial effect of authorizing the peaceful enjoyment of the right to copy and perform the work because the licensor do not actually own the rights she pretends to license: the absence of representation by the licensor that the work does not contain a copyright infringement, and the incompatibility of the system with collective management in case the licensor is a member of a collecting society which prevents her to exercise her rights individually (which will be studied in the coming sub-section d.)

Representations are a statement, an assurance to the other party, a declaration of facts. Representations address here the statement that the work does not constitute an infringement of third parties rights, namely a copyright infringement but potentially an infringement of other rights such as trademark, privacy, publicity, etc. Representations should be distinguished from warranty and liability addressing for instance the quality of the work seen as product available for sale and the fact that an educational or informational work does not contain factual mistakes or are fit to teach.

Version 3.0 clause mixes these different notions, while we address here only the absence of

representations or warranties concerning noninfringement:

5. Representations, Warranties and Disclaimer
UNLESS OTHERWISE MUTUALLY AGREED TO BY THE PARTIES IN WRITING, LICENSOR OFFERS THE WORK AS-IS AND MAKES NO REPRESENTATIONS OR WARRANTIES OF ANY KIND CONCERNING THE WORK, EXPRESS, IMPLIED, STATUTORY OR OTHERWISE, INCLUDING, WITHOUT LIMITATION, WARRANTIES OF TITLE, MERCHANTIBILITY, FITNESS FOR A PARTICULAR PURPOSE, NONINFRINGEMENT, OR THE ABSENCE OF LATENT OR OTHER DEFECTS, ACCURACY, OR THE PRESENCE OF ABSENCE OF ERRORS, WHETHER OR NOT DISCOVERABLE. SOME JURISDICTIONS DO NOT ALLOW THE EXCLUSION OF IMPLIED WARRANTIES, SO SUCH EXCLUSION MAY NOT APPLY TO YOU.

What is the point of using a CC work if it cannot be legally reused because the Licensee will after all not receive all the rights needed to use the work because the Licensor does not have them? We will discuss the pros and the cons of providing representations, and based the various clauses or absence thereof and analyze which options are viable for the legal validity of the system and the sustainability and the certainty of the downstream chain. Some consumer legislations forbid disclaiming certain warranties and some tort laws forbid misrepresentations.

A reason for removing the representation by the licensor that she holds the necessary rights to license them to the public between version 1.0 and version 2.0 was that it would not be fair to place the burden of due diligence and rights clearance on the licensor who already offering her work for free. An argument against representation by the licensor is the high damages that he might incur, at least in the US, where authors may be discouraged or prevented to distribute works if they have to carry the responsibility to check the status of every element of their work, especially without remuneration.

A specific use case in the mind of Creative Commons board members was documentaries which, except if they take place in an empty room with two of your family members, have a high risk of embedding copyrighted or otherwise protected elements. However, the 1.0 version warranty was not absolute but limited to the best of the knowledge of the licensor, and it is now one of the most dangerous caveats for the adoption of the system by professionals. Another argument for removing the representations from the license grant is that the warranty offered by an unidentified person who has two euros on her bank account would not be practically enforceable, while the work offered by a renowned institution would be. This observation relates to the identification of the parties: if the name of the Licensor is made available, it might provide a hint on the value of the grant.

The GNU-GPL and GFDL licenses, CC0 and Science Commons Protocol for Implementing Open Access Data, do not provide representation or warranty by the licensor that she has secured all the rights to permit the lawful and peaceful enjoyment of the rights granted by the license. Neither do they have a clause on representations nor do they expressly disclaim representation, meaning that rights ownership is a question of evidence which is left outside the contract, which is a reasonable middle ground between the two choices available in the CC licenses, limited representations or express disclaimer of representations.

CC licenses' initial version 1.0 and some jurisdictions versions, as well as the Free Art License[213] contain a limited representation and warranty by the author that the content does

[213] The freedom to use the work as defined by the Free Art License (right to copy, distribute, modify) implies that everyone is responsible for their own actions.

not infringe upon the rights of third parties. Also, the Public Domain Dedication included some representation[214].

Version 1.0 clause 5. entitled Representations, Warranties and Disclaimer would specify that the licensor owns the rights to secure a quiet use by the licensee: the licensor would warrant that the work does not infringe any rights, and that it can be used without paying royalties:

"By offering the Work for public release under this License, Licensor represents and warrants that, to the best of Licensor's knowledge after reasonable inquiry:
- Licensor has secured all rights in the Work necessary to grant the license rights hereunder and to permit the lawful exercise of the rights granted hereunder without You having any obligation to pay any royalties, compulsory license fees, residuals or any other payments;
- The Work does not infringe the copyright, trademark, publicity rights, common law rights or any other right of any third party or constitute defamation, invasion of privacy or other tortious injury to any third party."

This provision was favorable to the licensee and fostering reuse and remix. Its removal does not directly create incompatibility between works, but at an upper level is a big caveat for the sharing and remix culture. It prevents the peaceful enjoyment of CC works because it may be that CC works may not be used as offered in the license. In relation to the cascade of responsibility described in the 2.3.3 section, it is up to infringement procedures and contract law to decide whether a licensor who distributed a work for which she does not own all the rights (either because it contains someone else's work, or because she is a member of a collecting society and cannot offer a work free of charge for all the uses of the grant) can be held responsible if the grant is invalid and the right holder or the collecting society sues the licensee who was expecting to use a "clean" work.

The rational presented on CC blog is that warranties can be sold and that the sustainability of the ecosystem is turned into an optional business model: "licensors could sell warranties to risk-averse, high-exposure licensees interested in the due diligence paper trial, thereby creating nice CC business model."[215]

The absence of representation by the licensor transfers to the licensee the burden of risk assessment and rights holders' identification. The latter task is difficult if the licensor did not indicate her contact and may even be impossible to pursue in the absence of attribution notice as allowed by the protocol CC0. Besides, disclaiming responsibility for obtaining permission and waiving subsequent liability if works happen to be infringing third parties copyright may not be legal in some jurisdictions. Offering content with an uncertain legal status may be misleading for licensees who might be held liable for reusing content they thought they had the authorization to. It should be clarified who would be held liable in case of infringement, the licensor or the licensee, and what role community regulation and good faith may play compared to contractual and non-contractual liability (tort law).

This policy choice to stop offering a representation is at least irrelevant and may at most be leading to the invalidity of the contract, as warranties are mandatory in some jurisdictions and will apply regardless of a contradictory waiver. Here are a few examples based on general principles or extracted from specific pieces of legislation.

Good faith is an implicit principle of contract law and bad faith invalidates contracts.

---

[214] A certifier has taken reasonable steps to verify the copyright status of this work. Certifier recognizes that his good faith efforts may not shield him from liability if in fact the work certified is not in the public domain.
[215] http://creativecommons.org/weblog/entry/4216

Misrepresentations may lead a contract to be void and opens to remedies and damages.
Disclaiming responsibility for obtaining permission, offering, with an incitation to reuse, works which rights are not all cleared, and disclaiming liability is not legal in all jurisdictions. In France[216], a licensor is bound to offer peaceful enjoyment, therefore a contractual waiver is neither valid nor applicable, an author warrants in any case that she is the actual author of the work and that the work does not infringe third party's rights.
Product liability legislation[217] offers some answers to the question whether representations are compulsory if not implied. Special duties are imposed to professional suppliers of goods and services by contract and tort law.
According to the European Code of Contracts article 42[218], contracts limiting responsibility for *dol* and *faute grave* are void. According to the principles of European law regarding non-contractual liability (tort in common law), there is a duty to not give misleading information based on the Unfair Commercial Practices Directive[219], fraud remedies cannot be excluded.[220]

Even if it is difficult for a licensor to make the effort to secure every single piece of the work, it is important to raise awareness and if not re-incorporate a full and absolute representation, remove the waiver which is at most invalid and at least risking to make the system useless if licensee cannot rely on licensed works non-infringing nature.

d. Collecting societies

Another hidden difference between jurisdictions lays in the clause addressing collecting societies. Indeed, an important caveat of the licenses is that in most jurisdictions, collecting societies require their members to assign all their rights in present and future works. Thus, members cannot use a Creative Commons license, even for some of their works or some of their rights. Authors can license their non-commercial rights for free under a CC license, and assign the management of their commercial rights in theory in some collecting societies in some countries, primarily the United States, the Netherlands or Denmark. But collecting societies situation varies from country to country and users can't have the same level of expectations. This has been translated into a clause to signal that mandatory collective management in some countries in some cases does not conflict with the obligation to offer the work for free.

In the case of works which are licensed under a CC license while they should not, because their rightholders are a member of collecting society, but which are reused by a Licensee, such Licensee may commit involuntary infringement. Who will be held responsible and liable to the collecting society? The Licensee who acted in good faith or the Licensor who should not have used the license? The reasoning is also applicable to the usages of the work which

---

[216] Article 1626 of the French Civil Code.
[217] Duintier Tebbens Harry, *International Product Liability: A Study of Comparative and International Legal Aspects of Product Liability*, Alphen aan den Rijn: Sijthoff and Noordhof, 1979, 433 p.
[218] Gandolfi (ed.), *Code Européen des Contrats*, 2002.
[219] According to the Directive 2005/29/EC of the European Parliament and of the Council of 11 May 2005 concerning unfair business-to-consumer commercial practices in the internal market and amending Council Directive 84/450/EEC, Directives 97/7/EC, 98/27/EC and 2002/65/EC of the European Parliament and of the Council and Regulation (EC) No 2006/2004 of the European Parliament and of the Council ('Unfair Commercial Practices Directive')
[220] von Bar Christian, *Principles of European Law*. Volume 1, op cit, p. 495.

fall under compulsory collective management and can thus not be granted for free. And this information is only available in collecting societies status and national laws, which should be reflected in the license jurisdiction's version of the collecting societies clause, but will not be easily accessible to users of licenses of other jurisdictions.

5. Conclusion: options to mitigate risks and improve compatibility

This concluding section evaluates possible solutions to improve the infrastructure and prevent inconsistencies to jeopardize the licensing system. Some of these options are not desirable because they could bring more problems than they would solve, or impose a high burden on CC, while some propositions could be implemented easily. Some elements could be redrafted in the short-term without requiring much effort. Other more substantial points could evolve in the long-term, after more research and development on the user interface and the definition of community guidelines.

Based on conclusions reached at various stages of this study, proposed solutions to solve legal problems are mostly from logical and technical nature. I propose to improve the interface design, as well as to reorganize and redraft the text of the licenses in order to rationalize and simplify the whole system. The text of the licenses could also be shorter and in plain language, closer to a Commons Deed. There could even be only a single document merging the human-readable summary and the legal code. I also suggest stopping the legal porting process which introduces involuntary inconsistencies. Definitions would not be drafted according to any legislation. Instead of being localized into jurisdictions, the CC porting process could take place within user communities and focus on social governance rather than on legal normativity.

5.1 Improve the interface

5.1.1 Develop more technologies to better support the licenses requirements

Licenses are constituted by several layers linking to each other: a logo, a summary of the license, the legal text of the license and metadata. It is neither certain that licensees read the legal license, nor that all of them even notice the link which appears when the mouse is on the logo and then click on it. The embedding of one format inside of the other is an elegant and effective design, and the link to the license could also appear in a less hidden way to make sure everyone can take advantage of it, which is already the case in the notice text. Besides, the logo HTML code delivered when selecting a license and accompanied by a piece of text, the notice button, could contain more information, or provide fields to incentivize users to add more information, such as what item precisely constitutes the Work or who is the Licensor,
I put an emphasis on the important role of a fourth format in addition to the common deed, the legal code and the metadata: the button, which is often the only information a user will see. It contains the logo of the options and a link to the human-deed. The button has to be accompanied by a sentence, the notice, which is included in the HTML code delivered by the Choose your license interface: “This work is licensed under a Creative Commons Attribution 3.0 Unported License”. However, this notice is sometimes deleted by the users, or not expressed in a way which is specific enough. It could be customized to fit users’ needs, for instance to describe what is intended to constitute the “Work” to which the license is applied: "Copy the text below to your Web site to let your visitors know what license applies to your works", informs CC when providing the notice button text to be inserted on a website.

The absence of specification of what is actually licensed may impact the validity of the agreement. Here is the sentence used on the CC website: "Except where otherwise noted, content on this site is licensed under a Creative Commons Attribution 3.0 License". But this is not the sentence which is generated by the interface and it does not allow to formulate the sentence corresponding to the cases “where otherwise noted”. Therefore, further fine-tuning the sentence and transforming the word “Work” into one or more editable fields could raise awareness of the licensors and help make them specify what they intend to CC-license. An easy solution would be to propose for a few options (single work, general website) some easy to copy/paste HTML notice text. At a later stage or for more experienced users, it could be made explicit by the licensor what constitutes the *work* in the License Notice: the website as a whole, some of the individual works placed on the website, for instance only the text and the music but not the images, the music including lyrics, a composition and its performance and their fixation. Being specific is very important to know precisely what is being licensed and the inclusion of fields describing the work would ease that process. Indeed, it is not easy to figure out what constitutes a music composition.

Metadata have an underused potential. It should be more frequent to see licensors include additional information. Thus, the possibility to fill these fields could be expressed in a more assertive way, and the number of these fields could be increased:

- The format of the work (audio, video, text, image, interactive, other),
- The title of the work,
- The name of the author or entity the licensor wishes the licensee to attribute,
- The name and contact of the licensor, which are currently missing,
- “The URL users of the work should link to. For example, the work's page on the author's site.”,
- The URL of the source work if the work is derived from another work,
- A URL for more permission, where a user can obtain information about clearing rights that are not pre-cleared by your CC license.

This would make the licensing process longer, but more complete.

Automatic tagging tools can facilitate the respect of provisions which are often not respected by the licensee because the task is difficult to perform, such as attribution, license notice, and choice of options for derivatives.
The management of license requirements for derivatives can be improved by developing more technologies based on the ccREL. Extended information on attribution and modifications can be embedded into metadata which would follow the work during its lifecycle and be updated semi-automatically, for instance when saving or uploading a document or a wiki page, the software could prompt the author to fill attribution, URL and modification history fields.
When remixing two works licensed under different options, an expert system could easily prescribe the licensing options which can be chosen for the derivative work. This task could be operationalized through the metadata update process, when adding the name of the new author, the new URL, etc.

5.1.2 Remodel the acceptation infrastructure

In order to answer some issues raised by contract law, the infrastructure could be improved by

adding text or fields which would be editable by the licensor.

Following e-commerce and e-signatures legal framework, it could be suggested to introduce a click-wrap acceptation of the legal code for licensors, including future and CCi versions. This might improve the contracting process, but it would make the licensing process more cumbersome, so the option may not be desirable.

The question of the consent is taken into account by the PD certification where the licensor explicitly manifests and expresses her consent to the license by checking a box[221]:
"I have read and understand the terms and intended legal effect of this tool, and hereby voluntarily elect to apply it to this work."

CC0 also makes the licensor manifest her consent:
"I hereby waive all copyright and related or neighboring rights together with all associated claims and causes of action with respect to this work to the extent possible under the law."
"I have read and understand the terms and intended legal effect of CC0, and hereby voluntarily elect to apply it to this work."
A double-click confirmation is even required:
"Are you certain you wish to waive all rights to your work? Once these rights are waived, you cannot reclaim them."

If the name of the author is to be indicated because of the attribution requirement, there is no such obligation to include the contact of the *licensor* while that information is useful to ask for more permissions beyond the license grant for instance. It should be considered to include a field for the name of the licensor in the license, this could be achieved with values to edit, such as in the BSD license template.

The addition of a form similar to the CC Public Domain tools would solve both problems of consent regarding consumer law requirements, and lack of identification of the contact person, being author or licensor. CC Public Domain tools all require explicit consent from the licensor who is asked to provide more information than requested by the standard interface, such as the name of the author.

The Founders Copyright tool, which operates an actual rights transfer, makes the contractual process much more detailed: the licensor shall provide the name of the right holder. The question of rights representations is also addressed as the licensor is faced to a series of questions:
"Do you have exclusive rights to this work?
Are there parts of your work that are from other sources (quotes, pictures, etc.)?
Is this a derivative work? (includes translations)"

These questions could easily find a place in the standard acceptation infrastructure to secure the system and limit infringement, or at least to inform the licensor.

Similarly regarding the representation issue[222], the Sampling Choose your license interface carries a warning that the standard Choose your license interface could also display:
"Before you apply the Sampling License to your work, make sure you have the authority to license all the rights involved. Musical works, for example, often consist of multiple copyrights (composition, recording, lyrics)."

---

[221] http://creativecommons.org/choose/publicdomain-2

[222] Exclusion of representations and warranties of noninfringement and the limitation on liability for any damages are not legal in all jurisdictions.

5.1.3 Reverse the logic of the system

The licenses logic is structured around the elements BY, NC, ND, and SA. BY is not optional anymore, while the three other elements are the first information available to users in all situations:

- Both as a licensor selecting a license because choice is given among NC, ND and SA on the Choose your license interface,
- And as a licensee as the combination of the elements produces the name of the license and their initials are displayed in the logo.

However, the licenses are not limited to these three elements, which only modify a core grant composed of eight longer clauses. The core grant, common among all the licenses, is neither displayed in the Choose your license interface, nor expressed in the title of the licenses.
The core grant is the non-exclusive right to reproduce, perform and distribute the unmodified work for non commercial purposes, it contains many other clauses which are shared between all the licenses. Even if the optional elements are very important and modify to a great extend the core grant, their preeminence may contribute to hide the basic clauses of the licenses.
Instead of focusing on the choice of options which modify freedoms, why not invert the presentation and present freedoms as modified by options? It could be more logic to present first what is at the core of all licenses and will be modified by the choice of the licensor, instead of focusing on qualitatively crucial, but quantitatively minor elements.
Most of the text of the licenses is indeed the same in all the licenses, and this important part of the licenses is being hidden because of the prominent position of the optional elements in the most visible parts of the licensing process: the interface and the logo, which might be the only information read by those users who won't read the bottom of the Common Deed or the Legal Code.

The machine-readable code, or ccREL (Rights Expression Language) is an abstract model with the syntax and the semantic needed to describe copyright permissions and conditions and build automatized applications.
To improve the logic of the system, it could be considered to recraft the expression of the permissions in the human-readable layer with RDFa syntax, while making sure that current machine-readable expressions would still be supported.
In the interface, this change would be reflected in the license chooser which would present the core grant (copy, etc) before the optional elements, reflecting a positive ontology of the clauses, instead of not displaying prominently enough the core freedom and the core clauses.
In the human-readable layer, this would be displayed by even more illustrative icons and corresponding lines in the Common Deed (adding e.g. warranties, publicity rights, choice of jurisdiction if any…) Some additional icons have already been designed, coming from the GNU-GPL, GFDL and BSD CC wrappers [223] conditions (notice, source code, no endorsement) and could be reused in an extended human-readable illustrated format.

A positive logic expression would present first the core clauses offered by all the licenses, the right to share the work for non-commercial purposes only, with attribution and without modification (BY NC ND), which can be augmented by more freedoms by adding (SA) or

[223] http://creativecommons.org/licenses/GPL/2.0/
http://creativecommons.org/licenses/LGPL/2.1/
http://creativecommons.org/licenses/BSD/

removing (NC and ND) optional elements. It should first be assessed (by leading some experience based on the design ergonomy and the logic of names) if reversing the logic of the system could constitute a realistic and workable option at all, and if optional elements, instead of being expressed negatively (NC, ND), could be expressed as an addition: the right to share even for commercial purposes, the right to reuse and even to make modifications. It could then be determined whether SA constitutes a positive addition or a negative restriction to a core grant in order to implement a similar positive representation.

The user interface in the CC Lab[224] provides a powerful example of cognitive re-organization of the options around the core grant. Allowing to play with the license elements and aggregate them differently than on the usual license chooser interface[225] provides another visual representation of the positive rights[226] expressed by the main clauses from which the NC and ND options are taking away. This puzzle interface constitutes an interesting starting point for further research and testing on the logic of the system.
It has already been the case: the Commons deed http://creativecommons.org/licenses/by/1.0/ on February 1st 2004 was mentioning that the grant included the right "to make commercial use of the work". It is cognitively useful to also display the contrary of NC and the contrary of ND (commercial uses and derivatives allowed).

To sum up, the license elements are accessible before the main clauses they alter, and this in the license chooser interface, in the notice button and in the title of the license. The main clauses appear only in the Legal Deed, and to a lesser extend in the Commons Deed while it could be very informative to have them displayed earlier in the cognitive process of the user seeing a logo and an interface, then icons within a Commons Deed, and perhaps no Legal Deed.

Eventually, such a reorganization of the representation of rights and conditions among the core grant and the license elements could lead to another way to name the licenses. Title simplification is much needed, as the names of the licenses (both the acronyms within the logo and the extended name in the title) are too long and not necessarily meaningful to the average audience who often indicates incomplete information and declares that a work is licensed under a CC license without mentioning which one, if not under "the" CC license, while there is not only one license. However, changing them could be tricky.

---

[224] The user can play with the bricks of a license on the Freedoms License Generator available in the ccLab at http://labs.creativecommons.org/demos/freedomslicense/. This license engine is presented as a puzzle and may have different cognitive results on the understanding by the user than the usual license choser interface: "Not all combinations are possible, but as you experiment with the selections, you can see the different licenses that result."
[225] http://creativecommons.org/choose
[226] Towards the definition of a positive rights expression ontology, which could be then reflected in a new structure for the options, see Melanie Dulong de Rosnay, « An Action-Based Legal Model for Dynamic Digital Rights Expression », in Tom van Engers (ed), *Legal Knowledge and Information Systems*. JURIX 2006: The Nineteenth Annual Conference. Amsterdam: IOS Press, 2006, pp. 157-162. http://halshs.archives-ouvertes.fr/halshs-00120011

5.2 Simplify the system

5.2.1 Redraft the text of the licenses

Consumer law suggests to draft plain language licenses, avoiding legal language which is difficult to understand and hard to read.
An example of plain language licenses is provided by the legal code of New Zealand clustering rights under "You may", conditions under "You must" and restrictions under "You must not".
It is possible to go even further. The Commons deed and the legal code could be combined in a single short and human-readable document presenting all the clauses in the form of clustered bullet points drafted in non-legal language, illustrated by corresponding icons.
Another starting point is the document entitled "baseline rights" which identifies briefly and clearly most of the rights and conditions for both parties without hiding too much information[227]. It is addressed to the Licensor while the Commons Deed targets the Licensee, and it focuses on the core clauses, not on the optional elements

"All Creative Commons licenses have many important features in common.
Every license will help you
- retain your copyright
- announce that other people's fair use, first sale, and free expression rights are not affected by the license.

Every license requires licensees
- to get your permission to do any of the things you choose to restrict - e.g., make a commercial use, create a derivative work;
- to keep any copyright notice intact on all copies of your work;
- to link to your license from copies of the work;
- not to alter the terms of the license
- not to use technology to restrict other licensees' lawful uses of the work

Every license allows licensees, provided they live up to your conditions,
- to copy the work
- to distribute it
- to display or perform it publicly
- to make digital public performances of it (e.g., webcasting)
- to shift the work into another format as a verbatim copy

Every license
- applies worldwide
- lasts for the duration of the work's copyright
- is not revocable"

The Open licenses core freedoms and restrictions synthesis proposed in section 3.5.4 also provides a starting point towards a shorter text:

Freedoms: Rights to Use
An open license grants all the necessary rights to access, copy, perform, distribute and modify a work, including in a database, a collection or a modified version and all types of usage.
The work and its source should be legally and practically accessible and modifiable.

Admissible conditions: Credits, Notice and Metadata
The author may require the work to be accompanied in an unmodified way by:
- the name, URL or a link to the text of the license,
- the title of the work, attribution information (author, performer, other right holder, sponsor…) as well as

[227] http://wiki.creativecommons.org/Baseline_Rights

modification history of the work to the extend they are provided in a reasonable way according to standards of citation,
- digital signature, original source or location and other metadata.

Non acceptable restrictions: Legal, Technical and Economic Usage Restrictions
An open license should not accept or impose:
- legal restrictions on the exercise rights to limit the users who may exercise the freedoms or the territory, scope, domain or field of usage,
- technical restrictions to access, download and edit a digital copy of work (technical protection measure, compulsory registration, distribution in a non-copiable or non-editable format…),
- economic restrictions to access and copy a digital copy of the work (distribution for a fee, in a format which is not free of charge…)

But even before taking the important step to write only one short text, a reorganization of the legal code could improve the layout and the readability. It would be easy to reorganize and cluster thematics. Also, it could help to add subtitles inside the longest clauses, like in the Sampling licenses section 3, the Australian legal code section 3 and 4[228], or the New Zealand section 2[229], in order to improve their readability.

Following findings of section 2.2.3 analyzing the main clauses, starting with the Definitions, it shouldn't require much effort either to modify slightly the text in order to match international law definitions and include all notions (this, of course, not in the case definitions would stop being legal and ported, as suggested further in section 5.2.1). As discovered in section 2.2.3, harmonization of the notions covered in the licenses with concepts included in international conventions include:

- The first fixation of a film or broadcast in the definition of Work,
- All the elements of a complex Work (for instance, music composition, lyrics, performance and fixation for a recording)

Otherwise, Work could be defined simply as "the copyrightable of work of authorship and/or the other forms of creation protected by related rights".
Adaptations should include adaptation of a broadcast.
The definition and the difference between adaptation and collection raise issues. They are legal notions which are difficult to grasp for non-lawyers. An extension of the Share Alike clause and the disappearance of the Non Derivative clause (authorizing only Collections and not Adaptations) could make the difference between both legal concepts irrelevant. This would lead to decrease the number of licenses as well as to simplify the text and therefore avoid misunderstanding (for instance on the qualification of Adaptation when synching music on moving images even when using the music track in its entirety and without modification).

The license grant should include the rights of commercial rental and public lending.
The definition of Original Author should be clarified to avoid confusion between authors and rightholders.
The fair dealing clause could be entitled Limitations and Exceptions and specify that limitations to related rights, and not only limitations to copyright, are not preempted by the License's Restrictions and License Elements (for instance, that a performance can be parodied even if it is released under an ND license which reserves modifications).

The license grant clause should include related rights or other applicable rights. It could also

[228] http://creativecommons.org/licenses/by-nc-nd/3.0/au/legalcode
[229] http://creativecommons.org/licenses/by-nc-nd/3.0/nz/legalcode

be re-organized. The incompatibility with other exclusive agreements (such as underlined in the FAQs for rights assignments to collecting societies) or transfer of ownership and all exclusive rights through, for instance, a publication contract with an exclusivity clause, could be made clearer. The information is not hidden and is obvious for the specialist, but not for the layperson who is often not aware of notions such as:

- The meaning of exclusivity,
- The prerogative of the original right holder to exercise her exclusive rights, and
- The impossibility to grant exclusive rights to a collecting society or a publisher when using a CC license.

Thus, a clarification could avoid licensors the risk of committing to incompatible agreements and be unable to comply with both at the same time.
The collecting societies clause, as well as the NC clause, could well fit here for all licenses, while it is currently the case only the non-NC ones, and the following section could be renamed for instance Notices and Credit. The clause related to technical measures could also be relocated, same for the provision stating that rights can be exercised in all media and formats, and that technically necessary modifications are not considered to be Adaptations. These small modifications would improve the consistency of these complex texts which structure ends up being scattered.

In the Restriction section, it should be clarified whether the notice requirement provision also applies to uses arising from limitations to exclusive rights. The collective management clause, which is part of the restrictions for NC licenses and part of the license grant for non-NC licenses, could be more closely related to the royalty-free provision that it amends.

More substantial modifications could also be evaluated and considered.

Following the Share Alike clause validity and implementation difficulties discussed in section 2.2.3, the consequences of introducing sub-licensing should be further studied, as it would allow to have a direct relationship between each successive parties and then have B endorse some responsibility towards C and allow C to sue B if A sues C while B committed the infringement.

The moral rights clause could be redrafted, if at all maintained. One the one hand, the provision may impose more restrictions than the law, as publishers usually do not enjoy moral rights. One the other hand, the provision may exclude some parties from its scope while they benefit from such a protection; moral rights may exist for non-authors in some jurisdictions, for instance for performers and film-makers in Australia, the latter being producers, directors and screenwriters, the film-maker producer being not mentioned in the CC definition for Original Author (she can be included if considered as a creator). Therefore, it is recommended to change the clause accordingly, and create separated definitions (and a contact field to be filled by the Licensor when selecting her License) for Author and for the other Original Rightholders, in addition to Licensor who would be the current rightholder.
Other provisions of the licenses are related to the exercise of moral rights and reputation to a broader extend and could be placed together: the attribution clause, the right to not be attributed upon request of any Licensor on Collections and Adaptations, and the non-endorsement clause stating that attribution should not imply a support by the Original Author, the Licensor or the Attribution Parties.

The removal of the clause on limited representation of non-infringement causes

incompatibilities as some versions and jurisdictions contain such a clause while the Share Alike effect will remove the representation. There is no consensus on the need to provide such representations or on the necessity to not provide them. Instead of asserting or excluding representations, it could be considered to not mention them and leave the question outside the license to be decided through applicable law.

Last, the provision stating that a waiver of the terms of the license should be consented to in a written signed contract could be located closer to the provision allowing distributing the work under different conditions. It should be clarified whether the license constitutes the entire agreement because another concluded at a later stage may exist elsewhere and that the license may not be modified without the mutual written agreement of the Licensor and the Licensee. This language should be simplified.

The substantive content of some clauses could not only be clarified, made shorter and located elsewhere, it could be substantially simplified.

For instance, the Attribution clause is located in three subclauses which could easily be gathered, and it contains very specific requirements which go beyond legal and social norms, while it could be more limited.
In any case, in the absence of technologies to help understanding the potential of this clause and fields to be filled, it is doubtful that is exercised to its fullest extent by licensors and implemented to its fullest extent by licensees.
Making even more prominent the metadata fields in order to foster their use could help to actually get the following attribution information which is so often not made available, preventing to respect the requirement of carrying this information:
- The name of the author, licensor or any party,
- The title of the work,
- The source URL of the work, but also the source URL of the original work for derivatives,
- For derivatives, a credit identifying the original author, the use of the original work and changes which have been made.
This requirement is difficult to express. It could either be deleted, or transformed into a non-binding best practice in the line of section 5.2.3., or the sentence could be semi-automatically drafted in the spirit of section 5.1.1.

All these changes would lead to a more compact text and some decisions on how to handle or leave out of the license some issues such as representations, databases, the scope of the Share Alike clause and adaptations, moral rights, etc. should also be taken. Instead of having to consider all the legal and policy differences between licenses making it difficult to cross-license, dual-license or re-license works, this simplification would also ease the compatibility process with other open content licenses.

5.2.2 Options rationalization: generalization vs. customization

The number of options and core clauses can either be reduced or extended.

The Share Alike clause could come back to version 1.0 and require to license derivative work under only the same version, instead of also under a CCi version or a future version or a compatible license, which are per se different and raise the most problematic compatibility

issue. But the 2.0 update was a useful policy move and going backwards would drastically reduce remixing options and raise incompatibility among works.

It could be possible to simplify the system and offer fewer licenses, for instance stop offering the less popular licenses, or the licenses which do not offer sufficient freedoms (these two solutions are contradictory, as the NC option is widely chosen and a moderated approach of freedom contributed to the success of the licenses) or the option which creates uncertainty (NC again).
Providing only one license would certainly be difficult: the simplest BY, the copyleft BY-SA, the most popular BY-NC-SA, the most restrictive BY-NC-ND? Nevertheless, a definition of what constitutes freedom for non-software works, the field of CC, would be clearly beneficial. It would obviously limit one reason of incompatibility between works licensed under different options. It would reduce information costs for users who have to choose between different options. It would also decrease legal uncertainty if users do not fully realize to what combination of options they consent for licensors or are bound for licensees. If there was a stronger conceptual definition of what principles constitute freedom for CC, and fewer variations from that core, there will be fewer incompatibilities.

The opposite possibility could be to increase the number of options (e.g. add advertising, in order to specify and thus clarify the notion of NC). But this would lead to increased information costs and more incompatibilities among options. It might otherwise be advisable, instead of adding more options inside the licenses, to externalize some of them in the CC+ protocol: warranties and representations if they don't become standard again, parallel distribution clause if there is a use case, distribution of sources if they don't become standard, database rights if they are not to be already re-included in the related rights… Having additional clearly identified icons could answer to more needs but it would obviously not simplify the system.

Finally, two options may be considered to circumvent international law difficulties: introducing an international private law clause and removing localized clauses and ported licenses.

First, international private law principles led to consider introducing a private international law clause to designate applicable law and competent jurisdiction. It should be studied what happens without such a clause and how it could impact the porting porting process. Is it an option to introduce dual licensing according to principle of territoriality? To make differences visible outside the local legal dead or commented re-translations available? Or is the most viable option to stop the legal porting which adds complexity?

The most simple and effective solution to reduce both the number of licenses and international inconsistencies among jurisdictions' versions would be to simply stop the porting process, and offer only a translation of a revised generic/unported 4.0 version. Such a text, as described in previous section 5.1.3, would be drafted in plain English and could use sui generis definitions in order to rely neither on legal interpretation nor on any national or international legal definition of works, rights, etc which differ among legal systems. This solution has been chosen by the FSF for the GNU-GPL and the GFDL. Definitions are not based on any legal concepts, but on ad hoc vocabulary of the domain, translations are merely linguistic and do not have legal value.

Implementation issues in local jurisdictions with different, incompatible legislations would not disappear, but this problem is inherent to copyright law which is not harmonized, and solving it is not a responsibility CC can bear. Thus, not offering ported versions would lead to stop adding complexity and internal inconsistencies which threaten the validity of the assent for both the licensor who has expectations which may be disappointed, and for the licensee who may ignore to what conditions she consents, or consent to conditions which will change.

It is a good thing to propose linguistic translations to improve access, acceptability and understanding by non-native English speakers and this should not be interrupted. It was a good choice to implement the porting process in the first place, because it led to the structuration of local teams and the internationalization of a US-based project, including on the legal level. But it quickly became obvious that the legal porting would only be a minor task of the jurisdictions teams, who dedicate a much greater amount of time to explain the licenses, give presentation, discuss with stakeholders and users, as well as discuss implementation issues, perform research, develop projects and propose improvement of the licenses and their infrastructure in coordination with the other jurisdictions teams and the headquarters.
The porting process has been a useful constitutional moment for the development of the international network, but it raises too many legal issues to be maintained for the sole purpose of improving accessibility and enforceability in local jurisdictions, all the more than now the generic licenses are not based on US law anymore, as it was the case when the international porting process was developed.

5.2.3 Diminish the impact of the law

Coordination by external intermediaries and user communities could add much value to a simplified CC licensing text and infrastructure.

Formalities, registration and licensing metadata update for liability can be offered by third parties. Safe harbors for infringement by licensees and insurance mechanism and online dispute resolution mechanism could also be implemented by others than CC.

User communities or institutional entities (e.g. Wikipedia, universities, funders) could recommend or even mandate the use of only one of the licenses, as a top-down ideological prescription and after identification of the most appropriate license suiting particular needs. For instance, and this in addition of CC making options' features more accessible, they could explain that the Share Alike clause can reach a similar effect to the Non Commercial option as far as limiting commercial exploitation is concerned, and that reputation and integrity concerns leading to the choice of the Non Derivative options are already answered by the Attribution clause.

The CC porting process could take place not into jurisdictions, but within communities, relying on social governance to define implementation norms more than on legal normativity for enforcement. Best practices could be defined and implemented within certain creative or user communities: life science researchers, electronic musicians, non-profit broadcasters, commercial platforms, public libraries, collecting societies… Two topics provide a great experimentation and normative field to test such a practice: define Attribution and Non Commercial.

Norms vary among jurisdictions which apply national legislations, but also among user communities creating and enforcing social norms. A set of ethical principles described in an extended common deed or in a separated document may be more effective and accessible than a detailed doctrinal definition ported in a multiplicity of jurisdictions. Thus, instead of long binding licenses, or in addition to a shorter text, protocols and guidelines of "appropriate behavior"[230] developed by communities may still have a normative aspect, without legal uncertainty issues, and act as "conversational copyright" communication tools[231] rather than as mere legal contracts. The fact that there has been little case law so far may indicate that enforceability is difficult to reach by individual users, or that the licenses can be considered more as a communication tool than a legally binding and easily enforceable instrument. Both judges and users could use these soft law documents to better interpret and implement the licenses.

[230] See the norms for contributors and users of data developed by the Polar Information Commons community at http.www.polarcommons.org/ethics-and-norms-of-data-sharing.php which intends to regulate for instance attribution and notification.

[231] Carroll Michael W., "Creative Commons as Conversational Copyright", Villanova Law/Public Policy Research Paper No. 2007-8, in *Intellectual Property and Information Wealth: Issues and Practices in the Digital Age*, Yu Peter K. (ed.), Vol. 1, pp. 445-61, Praeger, 2007. http://ssrn.com/abstract=978813

6. References

Literature

ABELSON Hal, ADIDA Ben, LINKSVAYER Mike, YERGLER Nathan, ccREL: The Creative Commons Rights Expression Language, Communia First Workshop, Torino, January 2008. http://www.communia-project.eu/node/79

ALAI Memorandum on CC licenses, Columbia Journal of Law and the Arts, 29:3, 2006, p. 261-269.

ANGELOPOULOS Christina, "Creative Commons and Related Rights in Sound Recordings: Are the Two Systems Compatible?", Institute for Information Law, December 2009, 44 p. http://www.ivir.nl/publications/angelopoulos/Creative_Commons_and_Sound_Recordings.pdf

BOND Catherine, "Simplification and Consistency in Australian Public Rights Licences", (2007) 4:1 *SCRIPTed* 38.
http://www.law.ed.ac.uk/ahrc/script-ed/vol4-1/bond.asp
http://www.austlii.edu.au/au/other/AIPLRes/2006/21.html

BOURCIER Danièle, DULONG de ROSNAY Mélanie (eds.), *International Commons at the Digital Age - La création en partage*, Romillat, Paris, 2004, p.

BOYLE James, "The Second Enclosure Movement and the Construction of the Public Domain", *Duke Law and Contemporary Problems*, vol. 66, p. 33-74, 2003. http://ssrn.com/abstract=470983

CARROLL Michael W., "Creative Commons as Conversational Copyright", Villanova Law/Public Policy Research Paper No. 2007-8, in *Intellectual Property and Information Wealth: Issues and Practices in the Digital Age*, YU Peter K. (ed.), Vol. 1, pp. 445-61, Praeger, 2007. http://ssrn.com/abstract=978813

CHEN Shun-ling, "To Surpass or to Conform – What are Public Licenses For?", *University of Illinois Journal of Law, Technology & Policy*, Vol. 2009, Issue 1, p. 107-139.

CHELIOTIS Giorgos, "Creative Commons Statistics from the CC-Monitor Project", *presentation at the iCommons Summit*, Dubrovnik, June 14-17, 2007. http://creativecommons.org/weblog/entry/7551

COATES Jessica, "Playing Well With Others: Increasing Compatibility Between Commons Licences.", Workshop on Asia and Commons in the Information Age, Taipei, January 2008.
http://meeting.creativecommons.org.tw/program:playing-well-with-others

Dir. CORNU Gérard, *Vocabulaire Juridique Association Henri Capitant*, PUF Quadrige 4ème éd. 2003.

CREATIVE COMMONS, Defining "Noncommercial". A Study of How the Online Population Understands "Noncommercial Use", September 2009.
http://wiki.creativecommons.org/Defining_Noncommercial

DOBUSCH Leonhard, QUACK Sigrid, "Epistemic Communities and Social Movements: Transnational Dynamics in the Case of Creative Commons", MPIfG (Max-Planck Institut for the Study of Societies) Discussion Paper 08/8.
http://www.mpifg.de/pu/dp_abstracts/dp08-8.asp

DUINTIER TEBBENS Harry, *International Product Liability: A Study of Comparative and International Legal Aspects of Product Liability*, Alphen aan den Rijn: Sijthoff and Noordhof, 1979, 433 p.

DUSOLLIER Séverine, "The Master's Tools v. The Master's House: Creative Commons v. Copyright", *Columbia Journal of Law and the Arts*, vol. 29:3, Spring 2006, p. 271-293.

DUSOLLIER Séverine, "Sharing Access to Intellectual Property through Private Ordering", *Chicago-Kent Law Review*, 2007, p. 1391-1435.

DULONG de ROSNAY Melanie, « An Action-Based Legal Model for Dynamic Digital Rights Expression », in Tom van Engers (ed), *Legal Knowledge and Information Systems*. JURIX 2006: The Nineteenth Annual Conference. Amsterdam: IOS Press, 2006, p. 157-162.
http://halshs.archives-ouvertes.fr/halshs-00120011

DULONG de ROSNAY Melanie, "From Free Culture to Open Data: Technical Requirements for Access and Authorship", in Danièle Bourcier, Pompeu Casanovas, Mélanie Dulong de Rosnay, Catharina Maracke (eds.), *Intelligent Multimedia. Sharing Creative Works in a Digital World*, Series in Legal Information and Communication Technologies, Vol. 8, European Press Academic Publishing, Florence, 2010, p. 47-68.

ELKIN-KOREN Niva, "Creative Commons: A Skeptical View of a Worthy Pursuit", in HUGENHOLTZ P. Bernt & GUIBAULT Lucie (eds.), *The Future of the Public Domain*, Kluwer Law International, 2006. http://ssrn.com/abstract=885466

ELKIN-KOREN Niva, "What Contracts Cannot Do: The Limits of Private Ordering in Facilitiating A Creative Commons", *Fordham Law Review*, vol. 74, p. 375, November 2005.

FARCHY Joëlle, "Are free licences suitable for cultural works?", European Intellectual Property Review, 2009, vol. 31, n° 5, pp. 255-263.

GANDOLFI (ed.), *Code Européen des Contrats*, 2002.

GARLICK Mia. "Creative Humbug? Bah the Humbug, Let's Get Creative!" *INDICARE Monitor* 2, no. 5 (2005). http://www.indicare.org/tiki-read_article.php?articleId=124

GUADAMUZ Andres "The license/contract dichotomy in open licenses: a comparative

analysis", *University of La Verne Law Review* 30:2, 2009, pp. 101-116.

GUADAMUZ Andres, "Viral contracts or unenforceable documents? Contractual validity of copyleft licenses", *European Intellectual Property Review*, vol. 26 (8), 2004, pp. 331-339.

GUIBAULT Lucie, Creative Commons: Struggling to 'Keep it Simple', in E. Schweighofer & P. Sint (Ed.), Conference Proceedings KnowRight 08, Wenen: Österreichische Computer Gesellschaft 2008, p. 75-83.

GUIBAULT Lucie, van DAALEN Ot, "Unravelling the myth around open source licences : An analysis from a Dutch and European law perspective", Information technology & law series 8, The Hague: T.M.C. Asser Press 2006.

HARGITTAI Eszter, Skill Matters: The Role of User Savvy in Different Levels of Online Engagement, Berkman Luncheon Series, Harvard Law School, 23-06-2009.
http://cyber.law.harvard.edu/events/luncheon/2009/06/hargittai

HELLER Michael, *The Gridlock Economy: How Too Much Ownership Wrecks Markets, Stops Innovation, and Costs Lives*, Basic Books, New York, 2008, 304 p.

HELLER Michael, "The Tragedy of the Anticommons: Property in the Transition from Marx to Markets", *Harvard Law Review*, Volume 111 (3), January 1998, p. 621-688

HIETANEN Herkko, "A License or a Contract, Analyzing the Nature of Creative Commons Licenses", NIR Nordiskt Immateriellt Rättsskydd (Nordic Intellectual Property Law Review), 2007/6, 76, p. 516-535.
http://ssrn.com/abstract=1029366

ITO Joi, "Creative Commons: Enabling the next level of innovation", *What Matters*, McKinsey & Co, 30-10-2009.
http://whatmatters.mckinseydigital.com/internet/creative-commons-enabling-the-next-level-of-innovation.
Original unedited version: http://joi.ito.com/weblog/2009/10/30/innovation-in-o.html

KATZ Zachary, "Pitfalls of Open Licensing: An Analysis of Creative Commons Licensing", *IDEA – The Intellectual Property Law Review*, vol. 46, n°3, 2006, p. 391-413.

KENNEDY Charles H., Making Enforceable Online Contracts, *Computer law review international*, 2009, n°2, pp. 38-44.

KOELMAN Kamiel, "Waarom Creative Commons niet kan werken", *Computerrecht* 2009, p. 112.

LESSIG Lawrence, "Re-crafting a Public Domain", 18 *Yale Journal of Law & the Humanities* 56 (Special Issue 2006).

LESSIG Lawrence, *Free Culture - How Big Media Uses Technology and the Law to Lock Down Culture and Control Creativity*, The Penguin Press, 2004, 348 p.

LESSIG Lawrence, *Code and other laws of cyberspace*, Basic Books, New York, 1999, 297 p.

LINKSVAYER Mike, " Approved for Free Cultural Works ", *CC News*, February 20th, 2008. http://creativecommons.org/weblog/entry/8051

LINKSVAYER Mike, "CC and the Google Book Settlement", *CC blog*, 16-11-2009. http://creativecommons.org/weblog/entry/19210

MAKO-HILL Benjamin, "Towards a Standard of Freedom: Creative Commons and the Free Software Movement", July 2005. http://mako.cc/writing/toward_a_standard_of_freedom.html

MARACKE Catharina, "Creative Commons International. The International License Porting Project – origins, experiences, and challenges", in Danièle Bourcier, Pompeu Casanovas, Mélanie Dulong de Rosnay, Catharina Maracke (eds.), *Intelligent Multimedia. Sharing Creative Works in a Digital World*, Series in Legal Information and Communication Technologies, Vol. 8, European Press Academic Publishing, Florence, 2010, p. 67-88.

MERGES Robert, "A New Dynamism in the Public Domain", *The University of Chicago Law Review*, Vol. 71, pp. 183-203, 2004.

MÖLLER Eric, "The Case for Free Use: Reasons Not to Use a Creative Commons-NC License". http://freedomdefined.org/Licenses/NC and Creative Commons -NC Licenses Considered Harmful", *Kuro5hin*, September 2005. http://www.kuro5hin.org/story/2005/9/11/16331/0655

PALLAS LOREN Lydia, "Building a Reliable Semicommons of Creative Works: Enforcement of Creative Commons Licenses and Limited Abandonment of Copyright", *George Mason Law Review*, vol. 14, p. 271, 2007.

RADIN Margaret Jane, "Humans, Computers, and Binding Commitment", *Indiana Law Journal*, vol. 75, p. 1125-1161, 2000.

RICHARD Jones, EUAN Cameron, "Full fat, semi-skimmed or no milk today - creative commons licences and English folk music", *International Review of Law, Computers & Technology*, Volume 19, Issue 3 November 2005, p. 259-275.

ROSEN Lawrence, *Open Source Licensing: Software Freedom and Intellectual Property Law*, Prentice Hall, 2004, 432 p. http://www.rosenlaw.com/oslbook.htm

ST LAURENT Andrew, *Understanding Open Source and Free Software Licensing*, O'Reilly, 2004, 207 p.

SHAFFER VAN HOUWELING Molly, "The New Servitudes", *Georgetown Law Journal*, Vol. 96, p. 885, 2008. http://ssrn.com/abstract=1028947

SHAFFER VAN HOUWELING Molly, "Cultural Environmentalism and the Constructed Commons", *Law and Contemporary Problems* 70, Spring 2007, p. 23-50.

Study Group on a European Civil Code/Research Group on EC Private Law, von BAR Christian, CLIVE Eric (ed), *Principles, Definitions and Model Rules of European Private Law - Draft Common Frame of Reference (DCFR)*, Sellier, 2008, 396 p.

VÄLIMÄKI Mikko, HIETANEN Herkko, "The Challenges of Creative Commons Licensing", *Computer Law Review*, 6/2004 (Volume 5, Issue 6), pp. 172-177.

VODJDANI Isabelle, Le choix du Libre dans le supermarché du libre choix, 2004, 2007. http://www.transactiv-exe.org/article.php3?id_article=95

WEATHERALL Kimberlee, "Would you ever recommend a Creative Commons license?", Unlocking IP 2006 Conference, "Creating Commons: The Tasks ahead in Unlocking IP," UNSW AGSM, Sydney, 10-11 July 2006, Australasian Intellectual Property Law Resources (AIPLRes) 22, 2006.
http://www.austlii.edu.au/au/other/AIPLRes/2006/22.html

Cases

LG München I, Az. 21 O 6123/04 - Welte./.Sitecom Deutschland GmbH

US District Court Southern District of Indiana, vom 28.11.2005, 05-618 - Wallace v. FSF

LG Berlin, Az. 16 O 134/06 - WLAN-Router

LG Frankfurt a.M., Az. 2-6 O 224/06 - Welte./.D-Link Deutschland GmbH

LG München I, Az. 7 O 5245/07 - Welte./.Skype Technologies SA

US Court of Appeals for the Federal Circuit, Washington D.C., vom 13.08.2008, 2008-1001 - Jacobsen v. Katzer

BPatG, Az. 33 W (pat) 1/07 - "Open Source Broker"

Cour d'Appel de Paris, vom 16.09.2009, Az. 04/24298 - AFPA v. EDU 4

European Court of Justice, Directmedia Publishing GmbH v Albert-Ludwigs-Universität Freiburg, Case C-304/07), 9-10-2008.

Cour d'Appel de Paris, No 04/24298, Edu4 v. AFPA, 16-09-2009.
http://fsffrance.org/news/arret-ca-paris-16.09.2009.pdf

United States Court of Appeals for the Federal Circuit, No. 2008-1001, Robert Jacobsen v. Matthew Katzer, 13-08-2009.

Landgericht Frankfurt am Main, 21/09/2006. http://www.jbb.de/urteil_lg_frankfurt_gpl.pdf
LG Frankfurt/aM verurteilt Vertriebsgesellschaft wegen GPL-Verletzung

Landgerichts München I, 19/05/2004 No 21 O 6123/04
http://www.jbb.de/urteil_lg_muenchen_gpl.pdf

District Court of Amsterdam, Adam Curry v. Audax Publishing B.V., Case 334492/KG 06-176 SR, 9/03/2006; European Copyright and Design Reports, Sweet and Maxwell, Westlaw, septembre 2006, Curry v Audax Publishing BV, 2006 WL 2584400, [2006] E.C.D.R. 22 (RB (Amsterdam), Mar 09, 2006) (NO. 406921).

Lower court number six of Bajadoz, ordinary procedure 761/2005, ruling, nº 15/2.006, 2006 feb. 17, Sociedad General de Autores y Editores (SGAE) v. Ricardo Andres Utrera Fernández.

ProCD, Inc. v. Zeidenberg, 86 F.3d 1447 (7th Cir., 1996).

Kaufman v. American Express Travel Related Services, Inc., United States District Court, No. 07 C 1707, 2008 WL 687224 (March 7, 2008).

Specht v. Netscape Communications, 150 F. Supp. 2d 585 (S.D.N.Y.2001).

DE GRUYTER OPEN

Wroclaw Review of Law, Administration & Economics

THE NORMATIVITY OF THE PRINCIPLE OF MUTUAL TRUST BETWEEN EU MEMBER STATES WITHIN THE EMERGING EUROPEAN CRIMINAL AREA

*ALEKSANDRA SULIMA**

I. THE EUROPEAN CRIMINAL AREA – CONCEPT AND EVOLUTION

The European Criminal Area is the expression commonly used in legal doctrine and practise for the cooperation of the EU Member States within their criminal justice systems. This term is not derived from the EU Treaties, neither from the Treaty on European Union (hereinafter: TEU[1]) nor from the Treaty on the Functioning of the European Union (hereinafter: TFEU[2]). This term covers the complex structure of the various instruments and mechanisms functioning within the Member States. It includes the traditional international legal instruments of cooperation as well as innovative supranational methods[3]. However, it must be mentioned at the beginning that the EU has shifted its emphasis from the former method to the latter. Cooperation governed by the intergovernmental method always implies some political consensus during the conclusion of numerous bilateral and multilateral agreements, whereas the supranational method of cooperation requires more engagement and readiness not only from governments, but also from the organs and agencies which apply the respective law on a daily basis.

The advanced cooperation aiming at establishing a single judicial area in criminal justice within the European Union is of an unprecedented character. Under this type of cooperation, independent states decide to relinquish a part of their sovereign rights to a newly-established supranational body. The governments decide to delegate a portion of decision-making powers to a new authority. One may ask how it is possible to achieve such an advanced form of interstate cooperation between

DOI: 10.2478/wrlae-2013-0006

* PhD candidate, Department of International and European Law, Faculty of Law, Administration and Economics, Wroclaw University; ola.sulima@gmail.com

[1] Consolidated version of the Treaty on European Union [2012] OJ C326.

[2] Consolidated version of the Treaty on the Functioning of the European Union [2012] OJ C326.

[3] J Garstka, 'Współpraca w sprawach karnych' in F Jasiński, K Smoter (eds), *Obszar wolności, bezpieczeństwa i sprawiedliwości. Geneza, stan i perspektywy rozwoju* (Warsaw 2005) 341.

independent states. The justification for the concept of the European Criminal Area should be sought in the complex foundations of European integration. When analysing the foundations of European integration, we must take into consideration extensive historical, economic and socio-cultural factors. In this broad context one can understand how such an advanced form of cooperation within an area as delicate as criminal justice could be at all possible.

First of all, it must be remembered that European integration started a few years after the end of World War II. The European Coal and Steel Community, precursor of the present EU, began building the foundations of European integration in a post-war setting. The idea of closer cooperation in the respective economic sectors was an attempt to forestall similar global conflicts. What is more, subsequent historical events had a significant impact on the general international situation. These included the collapse of colonial empires and the constant military and economic threat of the Soviet Union. All of these factors shaped the reality of the 1950s, forming the political atmosphere within the countries of western Europe which decided to tighten what were initially economic relations.

Secondly, purely economic considerations cannot be forgotten. The commercial benefits resulting from cooperation between states located in one geographical region seemed at the time to be obvious.

Thirdly, there are also some factors enabling European integration of a socio-cultural nature. European countries share a cultural background and constitute a community of common values.

This comprehensive approach to the foundations of European integration allows us to understand the particular situation within EU Member States, the situation enabling the creation of more advanced mechanisms of cooperation. Analysis of this background leads us to the concept of mutual trust as a precondition of innovative cooperation[4].

At the same time, one cannot forget that historical and political dependences can sometimes constitute disintegrating factors, the so-called phenomenon of mistrust between the respective parties. However, it seems that the integrating factors within the emerging European Criminal Area are prevailing, that the concept of the EU viewed as a community of common values is well-founded and justified. Thus, it seems that cooperation between EU Member States in criminal matters based on the shared values of liberty, democracy, respect for human rights and fundamental freedoms, the rule of law and solidarity[5] has overwhelming potential to create a single area of justice in criminal matters. It is said that the EU has been able to create a community of mutual fate, a community guided by the idea of solidarity, which has placed human beings and their natural rights at the centre of its interests[6].

---

[4] M Ficher, *Mutual trust in European Criminal Law*, University of Edinburgh School of Law Working Paper Series (2009/10).

[5] A Grzelak, 'Wzmacnianie wzajemnego zaufania między państwami członkowskimi w obszarze współpracy w sprawach karnych jako czynnik integrujący Unię Europejską' in K Żukrowska (ed), *Co dzieli, co integruje Wspólnotę Europejską?* (Warsaw 2007) 411.

[6] K Popowicz, *Rozwój podstaw prawnych Unii Europejskiej*, vol. I (Warsaw 2009) 67.

II. THE CONCEPT OF MUTUAL TRUST

Many treatises elaborating the new area of cooperation of EU Member States in criminal matters use the term "mutual trust" as a basis and a precondition for the creation of a common area of criminal justice. They often refer to the example of the single market in economic cooperation as a model for the criminal area. The substance of this commonly used phrase, however, is not clear. One should ask what this concept includes, whether the term 'mutual trust' comes from social sciences and has no legal meaning, or is a term with extra-legal roots that has acquired a legal dimension.

Initial reflections on the essence of mutual trust lead to extra-legal associations. Thus, at the beginning of the analysis of the nature of the principle of mutual trust within EU structures, we should recall the basic assumptions of the sociological theory of trust. This sociological theory implies the creation of social ties and the formation of an institutional order based on social capital. This theory is crucial for further considerations of trust and the social foundations of public order. It is emphasized, however, that the sociological theory of trust is not sufficient to explain the complex political mechanisms in the sphere of internal security and justice developed within the EU. Assuming that the aim of EU Member States cooperating with EU institutions and bodies within the European Criminal Area is to strengthen the transnational supervisory and controlling mechanisms created for ensuring internal security, the sociological context does not fully illuminate the issue. However, there is no doubt that social expectations should be taken into consideration by the institutions responsible for public order in the functioning of such mechanisms. Of particular importance is the possibility of having trust in the legality and efficiency of law enforcement authorities, which play a fundamental role in the protection of citizens and society as a whole from pathological phenomena and threats[7]. Finally, it should be mentioned that one cannot exclude the usefulness of the sociological theory in the course of further analysis of the principle of mutual trust and its legal aspects, as this principle has not yet been precisely defined either in the Treaties or in Court of Justice jurisprudence[8].

Thus, considering that the sociological theory of trust is not alone sufficient to understand the European Criminal Area, the normative approach is invoked. This approach treats trust as a mechanism of cooperative actions based on the common norms[9]. The roots of this concept are to be found primarily in the general principle stipulated in Art. 4 (3) of TEU: "Pursuant to the principle of sincere cooperation, the Union and the Member States shall, in full mutual respect, assist each other in carrying out tasks which flow from the Treaties." The Court of Justice referred in its criminal cases to the principle of loyalty even before the communitarisation of the former third pillar[10].

---

[7] BA Misztal, *Trust in Modern Societies. The search for the bases of social order* (Cambridge 1996) chapter 1.
[8] Ficher (n 4) 12.
[9] F Fukuyama, *Zaufanie. Kapitał społeczny a droga do dobrobytu* (Warsaw 1997) 38.
[10] Ficher (n 4) 12

Focusing in particular on integration in the sphere of internal security and justice, one should remember the delicate nature of this area, which has always been considered under the dominion of national sovereignty. Over the fifty years of its development, the EU has created a multi-level security structure. This structure requires strong relations between the participants involved in creating this reality. The complex decision-making mechanisms, information flows and division of powers involved create what is at times not a very transparent system, the development of which is constantly in progress and involves continual coordinating efforts and stimulation. Thus, it is said that in order to make the system function smoothly, it is necessary to achieve operational efficiency, to jointly elaborate and implement effective policies, procedures and legal instruments as well as to assure advanced technological capacities in each Member State. To link the aforementioned elements of the emerging European Criminal Area we must also fulfil another condition, namely that of mutual trust between the interested parties. This trust must be the result of the political will of parties undertaking the joint initiative, and it must also reflect a readiness to enhance cooperation stemming from faith in the reliable and responsible approach of all participants to the agreed objectives and targets[11].

As far as mutual trust in cooperation regarding criminal matters is considered, both the subjective and objective scope can be recognized[12]. The former refers to a range of subjects invloved. It can concern state authorities as well as judicial authorities of other countries. This trust can be also be invoked in reference to vertical relations, namely between state authorities and individuals. This relationship must be interpreted broadly, going beyond the approach of state bodies to citizens. The idea of the European Criminal Area assumes that every individual should have the same high level of confidence in protection of the law within every Member State, irrespective of nationality[13]. Proceeding to an analysis of the objective scope of mutual trust, it must be said that this depends on the parties who are under consideration at a given moment, the nature of their interdependence and the type of cooperation. We can define the objective scope in many different ways, from general statements to very precise ones. Thus, trust can refer in general to the functioning of the national judiciary or the observance of the rule of law. Simultaneously, we can also speak of trust in sincere cooperation within certain legal instruments, e.g. trust in the reliable implementation of the Framework Decision on the European Arrest Warrant (EAW)[14] or in the due enforcement of a request for the surrender of a person.

The abovementioned examples of how mutual trust is required between Member States will not be analysed to an equal extent in this paper. In our consideration of the legal character of the principle of mutual

---

[11] A Gruszczak, *Współpraca policyjna w UE w wymiarze transgranicznym. Aspekty polityczne i prawne* (Wydawnictwo Uniwersytetu Jagiellońskiego 2009) 231-232.

[12] Ficher (n 4) 13.

[13] A Grzelak, 'Przestrzeń Wolności, Bezpieczeństwa i Sprawiedliwości' [2007] Sprawy Międzynarodowe 16.

[14] Council Framework Decision, (2002/584/JHA), on the European arrest warrant and the surrender procedures between Member States [2002] OJ L190/1.

trust, particular emphasis will be put on the relationship between the various national legislators – the authorities which create the law – and between the various executing bodies – the authorities which apply the law. As far as national legislators are concerned, the most common method used is that of intergovernmental cooperation – the traditional model of interstate actions. However, close cooperation at the level of the application of the law constitutes a novelty within international cooperation. The authorities which apply the law should have confidence not only in the legal systems of other Member States, but more pertinently, they should have confidence in the effects of the functioning of these systems, in the judgments and decisions issued by foreign bodies, as well as in the legitimization and competence of these organs to undertake certain actions[15]. All of these situations require each interested party to possess knowledge about other participants. Such mutual knowledge of participants' legal systems seems to be a key aspect in the process of building the European Criminal Area. Therefore, it seems that the essence of mutual trust is not the abstract belief that the other party shall comply with the common rules, nor is it the possession of pure, objective knowledge. Trust in fact constitutes a conviction that other Member States will comply with agreed-upon rules, this assumption being based on concrete, significant knowledge[16].

The statement that the crucial condition in the process of building mutual trust is increasing mutual knowledge should be complemented by the issue of the legitimization of actions[17]. Until the reform introduced by the Lisbon Treaty, the sphere of cooperation in criminal matters was highly criticized due to a lack of legitimacy and the lack of basic features required in democratic societies, as manifested by the weak role played by the European Parliament and national parliaments and the creation of law in isolation from the EU's citizens. The Lisbon reform has introduced, however, significant changes in this area, which can be deemed as a huge step towards reinforcing mutual trust in the emerging European Criminal Area.

To conclude these first considerations over the concept of mutual trust, it should also be emphasized that trust is, by its own nature, a dynamic phenomenon. Trust will always contain an element of risk. The process of building mutual trust between sovereign states is necessarily long, toilsome and requires the involvement of all parties. What is more, results already achieved can be lost very quickly. Thus, it has to be stressed that this process is continuous and will never be completed.

Furthermore, while one has to bear in mind this constant progress of the building of mutual trust, it also should be remembered that this process is not homogeneous. The level of already-achieved trust varies significantly in certain areas of cooperation, policy or even concrete legal instruments. The required, desired or existing degree of mutual trust between Member States may, thus, vary considerably and this should be considered when analysing every single legal institution[18].

---

[15] Ficher (n 4) 13.
[16] ibid 17-18.
[17] ibid 17.
[18] ibid 19.

III. THE CIRCUMSTANCES OF AN APPLICATION OF THE PRINCIPLE OF MUTUAL TRUST

At first sight it may seem that at each stage of cooperation between sovereign subjects the existence of mutual trust is necessary, that there is trust that another party acts in good faith, and that it can be relied upon to implement common standards. However, the necessity of the existence of mutual trust is not a uniform requirement for all forms of cooperation within the EU; its necessity depends on the one hand on the division of powers between the EU and the Member States, as well as the nature of these competences, while on the other hand it depends on the stage of cooperation, its advancement.

Firstly, the issue of EU competences should be mentioned. The sphere of the EU's competences can be divided into exclusive (Article 2 § 1 and 3 TFEU) and non-exclusive forms. The latter includes shared competences (Article 2 § 2 of the TFEU), which were called competitive competences prior to the Lisbon Treaty, and other competences to carry out actions to support, coordinate or supplement the actions of Member States without thereby superseding their competence in these areas (Article 2 § 5 of the TFEU), which were previously known as parallel competences.

The area of freedom, security and justice, and thus the emerging European Criminal Area as well, have been included in the group of so-called shared powers (Article 4 § 3 of the TFEU)[19]. The idea behind this type of competence is explained in Article 2 § 2 TFUE, which says that: 'When the Treaties confer on the Union a competence shared with the Member States in a specific area, the Union and the Member States may legislate and adopt legally binding acts in that area. The Member States shall exercise their competence to the extent that the Union has not exercised its competence. The Member States shall again exercise their competence to the extent that the Union has decided to cease exercising its competence.' This issue is also clarified in the Protocol on the exercise of shared competence[20]. The Member States declared that: 'when the Union has taken action in a certain area, the scope of this exercise of competence only covers those elements governed by the Union act in question and therefore does not cover the whole area.' The EU's competences within the area of freedom, security and justice therefore do not cover all issues in this field, but only those regulated by the basic act of the EU. The limitation of the EU's authority is also apparent through Article 72 TFEU, which guarantees the exclusive competence of Member States in the maintenance of public order and protecting internal security[21].

When considering the application of the principle of mutual trust, the relationship between the type of competences and the existence of mutual trust must first be analysed. If the Member States decide to grant the EU exclusive competence in a certain area, there is no further place for divagation as to whether and to what extent the necessity of building mutual

[19] Grzelak (n 13) 4, 133.
[20] Protocol No 25 to TEU, TFEU [2010] OJ C83/201.
[21] Article 72 (*ex* Article 64(1) TEC and *ex* Article 33 TEU) 'This Title shall not affect the exercise of the responsibilities incumbent upon Member States with regard to the maintenance of law and order and the safeguarding of internal security.'

trust is required; this is because the Member States, guided by a will for a certain policy to be effective at a higher level, have already decided to delegate some sovereign rights to supranational organizations. Nevertheless, the act of granting the exclusive powers itself certainly proves the existence of a high degree of trust between the participants of the cooperation.

As far as the principle of mutual trust is concerned, shared competences should be taken into consideration. This sphere of competence should also be analysed in the context of the principle of subsidiarity, which keeps the balance between the EU's intervention and the independent actions of Member States. The EU can intervene only if it is capable of acting more efficiently than the Member States. Thus, in the context of judicial and police cooperation in criminal matters as an area of the shared competences, there is a lot of space for Member States to undertake individual actions. The lack of exclusive competence of supranational organs causes a diversity; this diversity requires, in consequence, the necessity of strengthening mutual trust among all participants.

The application of the principle of mutual trust also depends on the particular moment of creation of the European Criminal Area. This can be illustrated by the example of the basic legal instrument which is used in criminal law cooperation, namely a directive. As far as the creation of common standards by means of directives is concerned, one can distinguish between three distinct stages. The first step is the creation of the EU law; in general this is made through an ordinary legislative procedure involving EU legislators, namely the Council and the European Parliament. The second step is the implementation of EU rules into national systems. The third step is the application of this legislation within individual cases. It is easy to observe that even if there is the political will to create a certain directive among the majority of participants, an individual Member State can destroy the aims of the respective law at the national level, e.g. through incorrect implementation. Furthermore, even with timely and correct implementation, a lack of trust among the authorities applying the law may lead to the complete lack of effectiveness of the EU law. Thus, it seems that mutual trust should be built with particular attention, especially between the organs which apply the law. It can be easier to conclude a pure declaration confirming the readiness for advanced cooperation and the necessity of building mutual trust, or even to adopt concrete measures via the Council and the European Parliament, the EU institutions where people are better acquainted with the EU's policies. The greatest challenge could be to build trust between the thousands of national officials and officers who apply the law every day.

IV. THE PRINCIPLE OF MUTUAL TRUST IN THE JURISPRUDENCE OF THE COURT OF JUSTICE OF THE EU

The principle of mutual trust is not mentioned in the Treaties. This does not, however, exclude analysis of the legal nature of this principle. In order to justify this approach, the fundamental principle of EU law can be evoked, namely the principle of the primacy of EU law. The Court of Justice

confirmed its existence in the 1960s, despite the fact that there had never been any specific provision alluding to this in the Treaties. In a pair of crucial judgements, i.e. the *Van Gend en Loos*[22] case in 1963 and *Costa v ENEL*[23] in 1964, the Court of Justice held that the Community had constituted a new, separate legal order where the Member States had limited their sovereign rights permanently and in consequence could not establish laws inconsistent with the essence of the Community. Moreover, the application of a law resulting from a treaty cannot be excluded by national law, because this would violate the community nature of the law and would undermine the legal basis for the functioning of the Community[24]. Although the principle of primacy is now unquestionable and has been confirmed by numerous decisions of the Court of Justice, even the last Lisbon reform failed to introduce the binding Treaty principle[25].

Therefore, the lack of a legal basis in the primary law of the EU does not preclude one from stating that mutual trust is not just a theoretical concept, but also has a normative context[26]. For confirmation of this thesis, a few significant decisions of the Court of Justice can be evoked here.

One of the most important judgements in criminal matters was issued in the case of *Gözütok and Brügge*[27]. It was the first judgment issued in the preliminary procedure in criminal matters as well as the first one regarding the interpretation of the Schengen acquis. The Court of Justice stated in the judgment's reasoning that the Contracting States have mutual trust in their criminal justice systems. Consequently, each of them recognizes the criminal law in force in other Member States, even if the application of its own national law would lead to another solution[28]. The Court of Justice confirmed the existence of mutual trust between EU countries, meaning that there is a conviction that the systems of justice in all Member States function efficiently. Moreover, at the same time the Court of Justice emphasized the existence of differences between countries, the fact of which, in its opinion, does not constitute an obstacle for mutual trust. In this particular case, the court ruled that the application by Member States of the principle of *ne bis in idem* from Article 54 of the CISA[29], in the context of proceedings leading to the expiry of the public prosecutor's right to investigate, which took place in another Member State without the participation of a court, cannot depend on the fact that the legal system of

---

[22] Case C-26/62 *Van Gend en Loos v Nederlandse Administratie der Belastingen* [1963] ECR I-00001.

[23] Case C-6/64 *Flaminio Costa v E.N.E.L.* [1964] ECR I-00585.

[24] P Mikłaszewicz, 'Zasada pierwszeństwa prawa wspólnotowego w krajowych porządkach prawnych według orzecznictwa ETS i Sądu Pierwszej Instancji, Omówienie wybranych orzeczeń 1963-2005' (Biuro Trybunału Konstytucyjnego, Zespół Orzecznictwa i Studiów 2005) 3, 4.

[25] ibid 1.

[26] A Grzelak, T Ostropolski, 'System prawa UE. Przestrzeń Wolności, Bezpieczeństwa i Sprawiedliwości UE. Współpraca policyjna i sądowa w sprawach karnych' in Jan Barcz (ed) vol. XI, part 1 (Warsaw 2009) XI.1-99.

[27] Joint cases C-385/01 *Gözütok* and *Brügge* [2003] ECR I-01345, I-1378.

[28] ibid section 33.

[29] Convention Implementing the Schengen Agreement, 14 June 1985, between the Governments of the States
of the Benelux Economic Union, the Federal Republic of Germany and the French Republic on the gradual abolition of checks at their common borders [2000] OJ L239.

the first Member State also does not require the participation of a court in such a case. Thus, if the system of one country does not have a certain legal institution, this state is obliged, based on mutual trust, to recognize the results of the application of the criminal law of another state. The Court stated that this is the only permissible interpretation of Article 54 which enables the effective application of this provision. This interpretation places emphasis on the object and purpose of the *ne bis in idem* principle, not procedural and purely formal aspects. Such an interpretation is necessary because of the diversity of legal systems within the EU's Member States; and indeed, this diversity should not be an obstacle to cooperation. Furthermore, thanks to the principles of mutual trust, it is not required to undertake a thorough harmonization of the rules in the criminal procedures of the Member States[30]. This broad interpretation of Art. 54 of the CISA in the context of mutual trust has become necessary for the effective application of this provision. This effectiveness has become one of the priorities of the EU, since the Treaty of Amsterdam incorporated the Schengen *acquis* into the EU *acquis*. Thus, the Schengen *acquis* is now one of the objectives within the development of the area of freedom, security and justice.

This ruling constitutes a clear and direct reference to the principle of mutual trust. However, it should be stressed that the Court of Justice only declared the existence of this phenomenon; it didn't create it. It should be noted that the Court's approach in this case is similar to the ruling regarding the issue of mutual recognition concerning one of the economic freedoms, namely the free movement of goods. The Court of Justice, in *Cassis de Dijon*[31], also invoked the principle of mutual trust between Member States as a basis for their duty to recognize the different standards that may be present in each state[32]. This judgment was the origin of the broad application of the principle of mutual recognition, as it stated that goods manufactured in accordance with the provisions of one Member State or introduced legally in a certain Member State's economic market should have access to other markets of EU Member States at the same conditions that exist in national ones. This general principle may be limited only in exceptional circumstances and cannot have a discriminatory character, so any such limitations must apply to national products in the same way as to foreign goods. Later, when the mutual recognition principle was already well established within the region's economic integration, it was also adopted in other areas of cooperation. Together with mutual recognition, a necessary precondition, namely mutual trust, has to be achieved. One can assume, thus, that the judgment which expanded and confirmed the application of the principle of mutual recognition and mutual trust for

[30]A Gajda, 'Trybunał Sprawiedliwości a III filar Unii Europejskiej' (2006) 2 Kwartalnik Prawa Publicznego 178.

[31] Case C-120/78 *Cassis de Dijon - Rewe-Zentrale AG v Bundesmonopolverwaltung fur Brantwein* [1979] ECR I-00649.

[32] N Thwaites, 'Mutual trust in Criminal Matters: the European Court of Justice gives a first interpretation of a provision of the Convention implementing the Schengen Agreement' (2003) 4(3) European and International Law 260.

cooperation within the former third pillar was that of the *Gözütok and Brügge* case.

What is of interest to legal scholars is the fact that this judgement is considered to be a form of indirect encouragement towards Member States to harmonize their criminal laws. Similarly, the abovementioned case of *Cassis de Dijon* prompted the Member States to agree common standards for the admittance of the goods in their markets. The Member States preferred to approximate their standards rather than to trust in different solutions and recognize them all at the same level. Even a partial harmonization can in fact establish minimum guarantees of the common principles[33]. In consequence, this has helped the Member States to treat the foreign goods in the same way as their national products. By way of analogy, we can say that within cooperation in criminal matters such minimal harmonization can help with the implementation of the *ne bis in idem* principle at the unprecedented, international context[34].

The significance of the *Gözütok and Brügge* judgment for the process of building the European Criminal Area is shown by the further practice of the Court of Justice. The Court has evoked the justification of this case very often, in many subsequent criminal cases. However, the revolutionary character of this decision has to be remembered. When this judgement was issued in 2003, criminal law cooperation was still subject to the intergovernmental regime of the third pillar. Thus, this courageous and decisive declaration of the Court concerning the necessity of the principle of mutual trust in so sensitive an area as criminal justice provoked many critical opinions. The main charge levied concerned the fact that the Court, a supranational institution, had started to be engaged intensively in matters where prior to that time it had played a very inconsiderable role. Thus, this ruling also became a catalyst for serious consideration of the communitarisation of the whole former third pillar[35].

This decision, as already mentioned, has been evoked in many cases within criminal matters. It is also worth discussing here another judgment of the Court of Justice issued during the preliminary ruling procedure in the criminal procedure case against *L.H. van Esbroesk*[36]. The Court once again interpreted Art. 54 of the CISA and repeated that the principle *ne bis in idem* means that there is mutual trust between the Contracting States with regard to their criminal justice systems. This is a significant decision as it included the phrase "the same acts" in the context of a provision Art. 54 of the CISA that says: ‘A person whose trial has been finally disposed of in one Contracting Party may not be prosecuted in another Contracting Party for the same acts (...).’ The Court found that the existence of different legal classifications for the same acts in the two countries cannot prevent the *ne bis in idem* principle from being applied. Due to mutual trust between Member States which compels them recognize certain activities as the same acts, the issue of legal classification is not important. The relevant criterion is the identity of material acts, irrespective of the legal classification or the

[33] ibid.
[34] H Hinterhofer, *European Criminal Law* (2011) 51.
[35] Thwaites (n 32) 262.
[36] Case C-436/04 *Criminal proceeding against L.H. van Esbroesk* [2006] ECR I-02333.

legal interest protected[37]. These material acts can be understood as the existence of inextricably linked behaviours. The effective application of the provision requires that the concrete wording of an offence cannot be relevant. It is natural that the legal systems of the EU Member States differ, but this is not an obstacle for the effective implementation of the *acquis* of the EU. Moreover, as has been already mentioned, the *ne bis in idem* principle in a unique, international context was incorporated into the EU *acquis*. It is worth noting here that the application of this principle in international cooperation is an innovative solution in the world and requires the huge commitment and confidence of the Contracting States. Though earlier conventions, such as the International Covenant on Civil and Political Rights[38] from 1966 and the 7th Protocol[39] attached to the ECHR, include *the ne bis idem* principle, it was only applied on the national level. Thus, these conventions declare the prohibition of instituting a new prosecution concerning the same crime under the jurisdiction of a single state. However, they don't prevent other states from initiating further proceedings. In fact, the Council of Europe adopted two conventions in the 1970s[40] introducing the international dimension of the *ne bis idem* principle, but they have not been ratified by most EU countries[41] and, therefore, they have only a marginal importance for the doctrine in a practice sense.

The next judgment which is worth evoking here was issued in the preliminary ruling procedure in the criminal case against *Dominic Wolzenburg*[42]. Advocate General Yves Bot, in his opinion attached to this case, stated that Member States, in agreeing to create a European Judicial Area in criminal matters and, in particular, the system of the European Arrest Warrant based on the principle of mutual recognition, waived a part of their sovereign powers. This renouncement means that they are not able to avoid the situation when the judicial authorities of other Member States institute investigations and prosecutions against their own citizens or to avoid the enforcement of sanctions issued by foreign authorities against their own citizens. The abovementioned Advocate General quoted the preamble of the Framework Decision on the European Arrest Warrant when explaining how this partial waiving of sovereign powers was at all possible. The framework decision provides that 'the mechanism of the European Arrest Warrant is based on a high level of confidence between Member States'[43]. The Advocate General concluded that the existence of mutual trust has been confirmed by several measures, *inter alia*, the withdrawal of the Member States from their right to prosecute in certain circumstances (the *ne bis in idem* principle as expressed in Art. 54 of the CISA). Yves Bot also cited the abovementioned decision of the Court of Justice issued in the joined cases of *Gözütok and Brügge.* He stressed that

[37] Hinterhofer (n 34) 54.
[38] International Covenant on Civil and Political Rights 1966, Art. 14 § 7.
[39] ECHR 1984, Art. 4 of Protocol No 7.
[40] European Convention on the International Validity of Criminal Judgments [1970] CETS No. 070, Art. 53, European Convention on the Transfer of Proceedings in Criminal Matters from [1972] CETS No. 073, Art. 35
[41] Hinterhofer (n 34) 51.
[42] Case C-123/08 *Criminal proceeding against Dominik Wolzenburg* [2009] ECR I-09621.
[43] Framework Decision of the Council 2002/584/EAW, section 10.

the *ne bis in idem* principle implies that, irrespective of which rules of sentencing are applied, each Member State has trust in their systems of criminal law and that each of them accepts a potential different outcome of the application of their laws. The analysed opinion concerning the case of *Dominic Wolzenburg* also provides us with a justification for the existence of mutual trust. He noted that trust among the EU Member States is based on many factors. First of all, each Member State, by acceding to the EU, was obliged to prove the observance of the fundamental rights defined in the European Convention on Human Rights and, since 7th December 2000, also defined in the EU Charter of Fundamental Rights. Moreover, the European Commission in the justification of the draft of the Framework Decision on the EAW, ensured that all Member States share a vision of the rule of law. On the other hand, the confidence which each Member State and its citizens should have in the other Member States' systems is based upon the logical and inevitable consequence of establishing the common market and European citizenship. It is worth noting here an interesting link between the functioning of the single market and the development of effective methods of the cooperation in criminal matters. On the one hand, such advanced economic cooperation can be a source of mutual trust within criminal cooperation. On the other hand, one of the fundamental provisions underpinning the European Criminal Area, namely Art. 54 of the CISA, is going to guarantee the free movement of EU citizens within the single market. The aim of this article is to ensure that no EU citizen that wants to exercise her/his right to free movement can be limited by the threat of consecutive prosecutions for the same acts in several Member States. Thus, both areas of cooperation, in economic and criminal matters, function on the reciprocity principle and they are linked together by strong interdependences.

Yves Bot also emphasized the particular role of mutual trust at the current stage of integration. He stated that due to the lack of thorough harmonization of procedural as well as substantive criminal law, the EU Member States, solely due to the functioning of the principle of mutual trust, can be convinced that the conditions under which their citizens would be prosecuted, tried and convicted in other Member States are at the same standards as the conditions in their own country. Thanks to the principle of mutual trust, the Member States can believe that other Member States will ensure that their citizens will be provided a thorough legal defence, regardless of the lack of knowledge of the language of certain proceedings or the rules of a specific procedure[44].

To end this analysis of the Court of Justice's approach to mutual trust, I would like to indicate that the national judiciaries of some Member States have also adopted this approach. First, it should be emphasized that the national courts are part of the judicial system within the EU. National judges, therefore, when ruling on matters subject to the EU's legal regime, become EU judges. An interesting judgment showing the influence of the Court of Justice on the national judiciary is the decision of the Polish Supreme Court in 2006[45] concerning the case of Adam G., who was

[44] Opinion of the Advocate General attached to case C-123/08 *Criminal proceeding against Dominik Wolzenburg* [2009] ECR 2009 I-09621, section 136.
[45] Resolution of the Polish Supreme Court (2006) I KZP 21/06.

prosecuted under an EAW issued by the Belgian authorities. In its resolution, the Polish court stated that the admissibility of a negative verification of the conditions under which a particular European Arrest Warrant was issued must be limited to very exceptional cases; it had been already forejudged by the obligation of mutual trust existing between the participants of European integration, the principle which constitutes the foundation of criminal cooperation among EU Member States.

V. THE LEGAL ASPECT OF MUTUAL TRUST

Turning now directly to considerations regarding the nature of the principle of mutual trust, one can start from the analysis of another case submitted to the Court of Justice in a preliminary ruling procedure, namely the case instituted by the German court in the course of criminal proceedings against *Klaus Bourquain*[46]. The opinion of Advocate General Dámaso Ruiz-Jarabo Colomer[47] is especially of importance here. He stated that though the principle of mutual trust is still novel in the whole concept of the European system of justice in criminal matters, it is however a foundation of the principle of mutual recognition, which undoubtedly constitutes the cornerstone of the emerging European Criminal Area. The Member States stressed this for the first time in the conclusions of the European Council from Tampere in 1999. The existence of mutual trust was also reported in the Council's Framework Decision on the European Arrest Warrant in 2002. In the preamble, it is expressly stated that the required high level of trust has already been achieved[48].

The Advocate General Damas Ruiz-Jarabo Colomer also noted that the reference to mutual trust in the abovementioned documents was a harbinger of the future significant, already discussed, Court of Justice's judgment in the joint case of *Gözütok and Brügge*. The Advocate General stated that the Court, at the first opportunity that it could, stressed the importance and obligatory character of mutual trust between Member States. The only possible broad interpretation of Art. 54 of the CISA is justified, in his opinion, by the principle of mutual trust. Thus, he claimed that mutual trust plays a utilitarian role, supporting the principle of mutual recognition[49]. Moreover, the *ne bis in idem* principle from the abovementioned article is inseparably linked to the requirement of the mutual trust, because regardless of whether the convergence of laws will become a reality someday, the effectiveness of Art. 54 of the CISA does not depend on the approximation of criminal laws of the Member States. He continued that the lack of the approximation can even be considered as an advantage because it compels the Member States to reinforce the scope of mutual trust, to spread common ground where trust can be built[50].

---

[46] Case C-297/07 *Staatsanwaltschaft Regensburg v. Klaus Bourquain* [2008] ECR I-09425.
[47] Opinion of the Advocate General attached to case C-297/07 *Staatsanwaltschaft Regensburg v. Klaus Bourquain* [2008] ECR I-09425.
[48] ibid section 39.
[49] ibid section 41.
[50] ibid section 44.

The Advocate General also raised the issue of the legal character of trust. He noted that, due to the fact there is no need for an approximation of laws, Member States are forced to accept that certain conditions, particularly in the field of fundamental rights, are being met by all participants of the integration. In this situation, he continued, mutual trust becomes a normative principle, one which combines the interpretative rules of the obligations resulting from the former third pillar. The principle of mutual trust thus plays a similar role to that of the principle of sincere cooperation, which has been already included in EU primary law[51].

One should also mention another judgment issued by the Court of Justice in the preliminary ruling procedure in the case of *Gasparini*[52]. The Dutch court asked the EU Court to determine whether the *ne bis in idem* principle, enshrined in Art. 54 of the CISA, applies to the decision of a court of a Contracting State which acquitted the accused due to the crime's period of prescription. The Court stated that in Art. 54 there is no reference to the particular content of the final judgment, thus there is no reason to interpret it in a restrictive way and limited its application only to convictions. The aim of Art. 54 of the CISA is the unlimited exercise of the freedom of movement by all EU citizens, including persons who were prosecuted and whose case was processed. The fear that new proceedings can be instituted against the same acts in another Member State is an inadmissible obstacle. At the same time, the Court of Justice noted that the degree of harmonization of criminal laws is insignificant, e.g. in this particular case, national laws concerning prescription periods are not comparable. However, no article of the EU Treaties or the Schengen Agreement or the CISA make the application of the Art. 54 dependent on such harmonization. In scholars' comments on the judgment, it was said that the Court of Justice took a maximalist approach to mutual trust, departing from the opinion of Advocate General Sharpston who stated, on the contrary, that the concept of mutual trust is not a reasonable basis for the broad application of the *ne bis in idem* principle in relation to ordering the dismissal of the prosecution because of a crime's limitation period[53].

VI. THE LIMITATION OF THE PRINCIPLE OF MUTUAL TRUST

One can ask whether advanced European integration requires unlimited trust from the Member States, as well as whether it is necessary or even desirable to have integration or a general crime prevention policy. As part of this paper's consideration of mutual trust, I would like to analyze the interdependences between the principle of mutual trust and the effectiveness of the emerging European Criminal Area.

In most scientific studies, scholars' opinions on a pro-integration, pro-EU approach can be found, emphasizing the effectiveness of new legal solutions. Thus, the restrained behaviour of some Member States is strongly criticized as a factor limiting the efficiency of criminal policies and the development of building the European Criminal Area. It is stressed that a

[51] ibid section 45.
[52] Case C-467/04 *Gasparini* [2006] ECR I-09199.

new dimension of advanced integration has also brought about new opportunities for the development of pathological phenomena, for the expansion of more dangerous criminal activity. Thus, the instruments which are to fight this phenomena also have to enter a new dimension of supranational cooperation. A European Criminal Area based on the principle of mutual recognition and mutual trust is considered the most effective answer for increasingly frequent global crimes as well as crimes resulting from the unique form of the integration and the openness of the Member States. However, there is also another side of the European Criminal Area and the functioning of its fundamental principles which are very often forgotten[54]: threats which can emerge once the European Criminal Area is fully realized. In this section, I would like to take into consideration a certain paradox, namely the link between strengthening cooperation and increasing the threat to the creation the area of freedom, security and justice. It should be noted that ill-considered integration between unprepared parties can badly affect crime prevention policies. One can enumerate, *inter alia*, the threat of bureaucratic opportunism, the abuse of fundamental rights and difficulties with taking responsibility for executing punishments at the European level[55].

However, one cannot argue that the EU seeks to strengthen cooperation between its Member States expecting mindless automatism, mere blind trust[56]. The judgment issued in the case of criminal proceedings against *Filomeno Mario Miraglia*[57] clearly shows that the Court of Justice put equivalent emphasis on the effective application of the principle of *ne bis in idem* based on the requirement of mutual trust, as it did on the efficient functioning of justice within the EU enabling the escape from criminal responsibility. Consequently, the Court of Justice found a balanced solution. On the one hand, the Court repeated the need for a broad interpretation of the *ne bis in idem* principle of Art. 54 of the CISA in order to avoid a situation where a person, through the freedom of movement, would be threatened by a subsequent prosecution for the same act in another Member State. On the other hand, the Court emphasized that this principle cannot be applicable to decisions issued without any examination of the merits. In the *Miraglia* case, the proceedings were ended by the prosecutor's decision only because of the previous institution of the criminal proceedings in another Member State against the same defendant for the same act. However, such a decision, according to the Court, cannot constitute a final judgment in the sense of Art. 54 of the CISA. The alternative interpretation would hamper the sanctioning of any punishable acts.

It may appear, in the practice of the functioning national authorities, that there is a risk of blind trust, of dangerous routine. The system established by the European Arrest Warrant based on a high level of trust can illustrate this threat. However, while the respective framework decision

---

[54] Ch Ecks, Th Konstadinides, *Crime within the Area of Freedom, Security and Justice. A European Public Order* (CUP 2012) 216.
[55] Z Brodecki, *Europa sędziów* (LexisNexis 2001) 26.
[56] Grzelak, Ostropolski (n 26) XI.1-120.
[57] Case C-469/03 *Criminal proceeding against Filomeno Mario Miraglia* [2005] ECR I-0200.

requires the proper control of each issued EAW in every situation, it does not always prevent mistakes due to automatism. To demonstrate this, the case of *Praczijk* can be evoked. He was a Belgian who was surrendered by his national authority at the request of the Italian organ. When he was arrested and sent to Italy, it was found that there was a mistake and the Italian authorities released him. The Belgian authorities stated that were not responsible for the mistake and they were not obliged to pay any compensation, evoking as the justification the principle and duty of mutual trust[58].

Therefore, it is said that the observance of the principle of mutual trust does not require absolute confidence. This trust can be defined as conditional[59], meaning that we can only speak of possible mutual trust within the EU thanks to many integrative factors such as the relative homogeneity of the standards of European legal systems, the identity of fundamental principles, and common interests arising from the geographical location of states. However, the existence and extent of this trust is not finally determined and it depends on further action to be undertaken by the interested parties[60].

VII. THE PROBLEM OF VIOLATION OF MUTUAL TRUST BY AN EU MEMBER STATE

The effects of a lack of trust within cooperation can be manifested at the level of the creation of the law as well as its application. First, the lack of confidence can be shown by the incorrect, limiting implementation of various legal instruments or significant delays in the transposition of the EU acts to national legislation. Secondly, the over-controlling approach of the enforcing organs to the legal instruments coming from other Member States or the extensive interpretation of the exclusion from mutual recognition can hamper the whole process of cooperation.

The consequences of lack of trust can be analyzed in two different situations. On the one hand, it should be considered in the context of further cooperation between states. The Member States often refer to the principle of reciprocity. Thus, a breach of trust by one party leads, in general, to the same behaviour of the other, even though it is stated in doctrine that the source of the EU Member States' obligation to fulfil the agreed tasks can no longer be derived from the reciprocity principle. The EU has already entered into a more advanced stage of integration and the traditional rule of reciprocal behaviours has already lost relevance. The source of EU Member States' obligations is now the common aspiration to remain loyal to the agreed obligations and to ensure that the European Criminal Area materializes. Therefore, of more relevance is the principle of sincere cooperation from Art. 4§3 TEU which, in the opinion of some scholars, has replaced the traditional rule followed in international relations, namely the

[58] A Weyembergh, 'Wzajemne uznawanie orzeczeń w sprawach karnych – bilans Europejskiej Przestrzeni Sądowej' in A Frąckowiak–Adamska, R Grzeszczak, *Europejska Przestrzeń Sądowa* (Wydawnictwo Uniwersytetu Wrocławskiego 2010) 56.
[59] Ficher (n 4) 13.
[60] ibid 13.

reciprocity principle[61]. However, some examples have shown that reciprocity is enshrined very deeply within international cooperation, and also remains so between EU Member States. Additionally, even though reciprocity is to be substituted, there are still some remnants of this traditional principle. The problem with the implementation of the EAW in Germany shows this clearly. The German Constitutional Court found[62] that the legislation transposing the EAW is unconstitutional and invalidated it[63]. From that moment until the entry into force of the new law, Germany did not adhere to the EAW system. In such circumstances, some Member States refused to act upon German warrants, referring directly to the principle of reciprocity. They claimed that if Germany did not comply with the principle of mutual trust, it cannot be expected that other countries will surrender their citizens[64]. Thus, the situation blocked cooperation in this particular field, but it could also have had more far-reaching consequences for the development of the whole criminal area, undermining the credibility of one of the parties.

The lack of trust can be also manifested in the way laws are implemented. One can take into consideration the example of the Polish legislation transposing the Framework Decision on the EAW. It should be noted that the first attempt of the Polish legislator, similar to the German approach, was recognized by the Constitutional Tribunal as unconstitutional[65]. The Polish Constitution, in the former Art. 55 § 1, stated that the extradition of a Polish citizen is prohibited. The legislator tried to justify its proposal, saying that the EAW was a completely different form of cooperation than extradition. However, the Constitutional Tribunal concluded that both instruments, the traditional extradition as well as the EAW, have an identical essence. Ultimately, the legislator was forced to change the constitution. The new version of Article 55 of the Polish Constitution[66] enables a Polish citizen to be surrendered under certain conditions, generally including the dual criminality requirement. However, it is said that such a solution does not provide full compliance with the Framework Decision on the EAW. It allows a departure from the verification of double criminality only in cases when a respective person is a foreigner. The framework decision differentiates situations when the dual criminality requirement exists or not, based only on the type of crime, not the nationality of the prosecuted person[67].

It should also be noted that there are further doubts with regard to the Polish implementation of the decision which relate to the provisions concerning the possible refusal to enforce specific EAWs. The directive provides three obligatory situations, whereas the Polish Code of Criminal Procedure[68] includes six and one of them relates to political crime (Art. 607p par.1 sec. 6); the EU legislation liquidated the political condition. It

[61] A Grzelak, M Królikowski, A Sakowicz, *Europejskie Prawo Karne* (CH Beck 2012) 58.
[62] Judgement of the Federal Constitutional Tribunal of 18.05.2005.
[63] H Saztger, *International and European Criminal Law* (CH Beck 2012) 121.
[64] H Kuczyńska, *Wspólny obszar postępowania karnego w prawie UE* (Scholar 2008) 164.
[65] Judgement of the Polish Constitutional Tribunal of 27.04.2005, P I/05
[66] Official Journal of Poland (2006) No. 200, position 1471.
[67] Grzelak, Ostropolski (n 26) XI.1-108.
[68] Polish Code of Criminal Procedure [2013] OJ 1529, item 480.

seems that the Polish implementation remains one step behind in comparison to the newest EU solutions. What is more, the Polish Code adds a second paragraph to Art. 607p, where the criterion of the citizenships is included. This addition comprises another obligatory condition in the situation where a respective person is Polish and the crime was committed on Polish territory. The directive provides territory rules only within the facultative conditions of refusal and does not make it dependent on citizenship. One can draw the conclusion that such differences in national legislations, the lack of the full conformity with the EU directive and the lack of consistency within all 28 Member States' solutions can lead to mistrust, to the undesired differentiation of the situations of EU citizens.

CONCLUSION

Based on the Court of Justice's jurisprudence and the practice of Member States, one can assume that a normative character of the principle of mutual trust has emerged. However, it cannot be allowed to prevent the continuous development of the foundations of this trust between Member States. The normative character of the principle enables Member States to better execute their obligations. However, the principle of mutual trust should not become a static rule. On the contrary, it should be the subject of constant progress. Recognition of the elaborated normative character cannot be synonymous with the abandonment of efforts by the EU and its Member States to build and strengthen the foundations of mutual trust. Normativity of the principle, along with realization in practice, can lead to the full creation of the European Criminal Area. It should also be emphasized that the recognition of the normative character of trust within the EU can have significant importance in other contexts, especially in the diversity of the process of building mutual trust and the lack of homogeneity in this area. It has already been stated that the level of trust currently achieved varies significantly in certain areas of cooperation, policy or even concrete legal instruments. In such a situation, the normativity of this principle shall guarantee equal standards within the whole European Criminal Area.

Wroclaw Review of Law, Administration & Economics

Journal Details

The Normativity of The Principle of Mutual Trust Between EU Member States within the Emerging European Criminal Area

Aleksandra Sulima

Published Online: 29 May 2014

Page range: 72 - 89

DOI: https://doi.org/10.2478/wrlae-2013-0006

THE NORMATIVITY OF THEPRINCIPLE OF MUTUAL TRUST BETWEEN EU MEMBER STATES WITHIN THE EMERGING EUROPEAN CRIMINAL AREA

ALEKSANDRA SULIMA*

I. THE EUROPEAN CRIMINAL AREA – CONCEPT AND EVOLUTION

The European Criminal Area is the expression commonly used in legal doctrine and practise for the cooperation of the EU Member States within their criminal justice systems. This term is not derived from the EU Treaties, neither from the Treaty on European Union (hereinafter: TEU[1]) nor from the Treaty on the Functioning of the European Union (hereinafter: TFEU[2]). This term covers the complex structure of the various instruments and mechanisms functioning within the Member States. It includes the traditional international legal instruments of cooperation as well as innovative supranational methods[3]. However, it must be mentioned at the beginning that the EU has shifted its emphasis from the former method to the latter. Cooperation governed by the intergovernmental method always implies some political consensus during the conclusion of numerous bilateral and multilateral agreements, whereas the supranational method of cooperation requires more engagement and readiness not only from governments, but also from the organs and agencies which apply the respective law on a daily basis.

The advanced cooperation aiming at establishing a single judicial area in criminal justice within the European Union is of an unprecedented character. Under this type of cooperation, independent states decide to relinquish a part of their sovereign rights to a newly-established supranational body. The governments decide to delegate a portion of decision-making powers to a new authority. One may ask how it is possible to achieve such an advanced form of interstate cooperation between

DOI: 10.2478/wrlae-2013-0006

* PhD candidate, Department of International and European Law, Faculty of Law, Administration and Economics, Wroclaw University; ola.sulima@gmail.com

[1] Consolidated version of the Treaty on European Union [2012] OJ C326.

[2] Consolidated version of the Treaty on the Functioning of the European Union [2012] OJ C326.

[3] J Garstka, 'Współpraca w sprawach karnych' in F Jasiński, K Smoter (eds), Obszar wolności, bezpieczeństwa i sprawiedliwości. Geneza, stan i perspektywy rozwoju (Warsaw 2005) 341.

J Value Inquiry (2015) 49:543–550
DOI 10.1007/s10790-015-9524-7

Ritual and Ritual Obligations: Perspectives on Normativity from Classical China

Michael Puett[1]

Published online: 19 September 2015

The goal of this paper is to explore some of the theories that arose in classical China concerning the ways in which normativity could be construed in ritual terms. I focus particular attention on the theories that were developed in portions of the *Book of Rites* (*Liji*), which would become one of the most influential bodies of ritual theory throughout East Asia. I will argue that some of these theories deserve to be incorporated into our contemporary philosophical understandings of normativity.

To make this argument, I will be turning to some rather – given my ultimate goals in this paper – counterintuitive material, including early Chinese discussions of mourning rituals. But I will try to argue that this material has something to offer contemporary discussions.

1 Rituals and Norms

Let me begin by mentioning one of the common criticisms made against the argument that we should take theorists of ritual seriously when discussing norms. Perhaps one of the reasons we often denigrate the significance of rituals is that we think rituals essentially function to socialize us into a way of thinking and acting. In other words, rituals tell us what to do. So, if the norms embedded within a given ritual are good, then the ritual would perhaps be acceptable. But what if the norms are not, from our point of view, acceptable? Following them would then entail submitting ourselves to "traditional" values, as opposed to having norms that could at least potentially be rationally adjudicated and altered.

It is, of course, entirely possible to have ritual traditions that become rigidified and function in a top-down way to instill values on a population. But thinking that

✉ Michael Puett
puett@fas.harvard.edu

[1] Harvard University, Cambridge, USA

this is the only way rituals can work is very limiting. It may therefore be worthwhile to look at indigenous theories of ritual from a so-called traditional society – in this case, classical China – to see how these issues are conceptualized. Intriguingly, they offer not only a suggestive way to think about rituals, but also a suggestive way to think about norms in general.

2 The Work of Mourning Rites

The chapters of the *Book of Rites* (*Liji* 禮記) were written over the course of the fourth, third, and second centuries BCE, and were compiled into a single text by the first century BCE. The text was thereafter defined as one of the "Five Classics," and became part of the standard educational curriculum throughout East Asia for much of the subsequent two millennia.

One of the key ideas underlying these chapters is the notion that ritual should transform our emotional dispositions.[1] In any situation, humans have tendencies to respond to situations in particular ways, and the goal of rituals is to shift these responses. For example, when someone passes away, we tend to be overtaken with grief and even anger. Mourning rituals force a change.[2]

The "Tan gong" chapter of the *Book of Rites* takes the reader through the early stages of the rituals immediately following a death. The chapter argues that the ritual world of mourning helps to modulate one's grief over a death and to alter those emotions such that they are shifted into a sense of remembrance of that from which one came:

> The rites of mourning are the extreme [expression] of grief and sadness. In modulating grief, one [learns to] accord with the changes [of life and death]. This is how the refined person remembers from where he came.[3]

To ensure this transition, the practitioner is taken through a series of steps in the ritual.

At first, the mourner expresses his grief toward the deceased:

> Bowing and hitting one's head on the floor is the extreme pain of grief and sadness. Hitting one's head on the floor is the depth of the pain.[4]

[1] For a fuller discussion of ritual theory in the *Book of Rites*, see Michael Puett, "Ritual Disjunctions: Ghosts, Philosophy, and Anthropology," in Veena Das, Michael Jackson, Arthur Kleinman, and Bhrigupati Singh, eds., *The Ground Between: Anthropologists Engage Philosophy* (Durham: Duke University Press, 2014), pp. 218–233; and Michael Puett, "The Haunted World of Humanity: Ritual Theory from Early China," in J. Michelle Molina and Donald K. Swearer, with Susan Lloyd McGarry, eds., *Rethinking the Human* (Cambridge: Center for the Study of World Religions, 2010), pp. 95–111.

[2] For a fuller discussion of mourning rituals in the *Book of Rites*, see Michael Puett, "Combining the Ghosts and Spirits, Centering the Realm: Mortuary Ritual and Political Organization in the Ritual Compendia of Early China," in John Lagerwey and Marc Kalinowski, eds., *Early Chinese Religion: Shang Through Han (1250 BC–220 AD)* (Leiden: Brill, 2009), pp. 695–720.

[3] *Liji*, "Tan gong," Chinese University of Hong Kong, Institute of Chinese Studies, Ancient Chinese Text Concordance Series (hereafter cited as ICS), 23/4.15/11. My translations here and throughout have been aided greatly by those of James Legge, *Li Chi: Book of Rites* (Oxford: Oxford University Press, 1885).

[4] *Liji*, "Tan gong," ICS, 23/4.15/12–13.

The grief is played out ritually by bowing to the deceased and hitting one's head on the floor.

After the deceased is interred in a tomb, the mourner returns home and wails. This is the point of most extreme grief, as the deceased has now left the home forever. The mourner is consoled by the rest of the mourning party upon returning:

> Consoling when [the descendant] returns [from the tomb] wailing is because this is the extremity of grief. He returns and there is no one there; he has lost [the deceased]. Therefore it is the most intense.[5]

The wailing upon the return is part of the ritual. The wailing at this moment is defined as being at its strongest.

And it then has to stop. Following the internment, sacrifices to the deceased are begun. A transition must now be made from mourning the deceased to sacrificing to the deceased as an ancestor. All wailing must cease at this point:

> The end of wailing is called "completing the event." On this day, auspicious sacrifices replace sacrifices of mourning. The next day, [the tablet] is enshrined with the grandfather.[6]

An ancestral tablet is made for the deceased, and the tablet is enshrined in the ancestral temple. The spirit of the deceased will now be called to the tablet. The enshrining of the tablet must be done quickly – the living cannot bear the idea of the spirit not having a place to return, now that the corpse has been buried in a tomb.

> Changing to auspicious sacrifices, and on the succeeding day to the enshrining of the tablet, must necessarily occur very close to this day. He [the survivor] cannot bear one day without a place [for the spirit] to return.[7]

The transition must now be complete from the mourning period to the subsequent period in which the deceased will be worshipped as an ancestor. The crying must end, and the name of the deceased must no longer be mentioned. He is now an ancestor, and will be worshipped as such, with an impersonator playing the role of the ancestor in accepting the sacrifices.

> One performs the sacrifice of repose and sets up the impersonator. There is a bench and a mat. One brings to an end the crying and avoids [the name of the deceased]. The services for him as living are stopped and the services for the ghost begin.[8]

At this point, the goal is to stop grieving and start developing properly filial dispositions toward the deceased as an ancestor.

Elsewhere in the *Book of Rites* the distinction is made clearly:

[5] *Liji*, "Tan gong," ICS, 23/4.15/20.

[6] *Liji*, "Tan gong," ICS, 23/4.15/21–24.

[7] *Liji*, "Tan gong," ICS, 23/4.15/24–25.

[8] *Liji*, "Tan gong," ICS, 28/4.52/6.

> In sacrificing, one is called "filial son" and "filial grandson." In mourning, one was called "grieving son" and "grieving grandson."[9]

The arc of the ritual thus takes the living from one of mourning for the deceased to one of sacrificing to the deceased in the form of an ancestor. If the ritual is effective, it means that the dispositions have been altered in a direction deemed better for the living.

This shift in the ritual from grieving for the deceased loved one to treating him as an ancestor itself operates entirely in the ritual space. Needless to say, there is no such clean shift in our emotional dispositions. After the ritual ending of the crying and the ritual beginning of ancestral sacrifices, one will of course continue to cry outside of the ritual space. But, if the rituals are effective, they create a space of practice within which different dispositions are being developed – dispositions that will be in tension with our sense of loss from the death of a loved one.

Once the transformation to ancestral sacrifices occurs in the ritual, a bifurcation develops for the living. Within the ritual space, the practitioner is being trained to think of the deceased as an ancestor and to repeatedly inculcate within himself the sensibility of remembering from where he came. But this is not a simple norm that will guide his life. Indeed, the reason the ritual has to be constantly repeated is to re-inculcate these dispositions over and over, as they are hardly the values that are guiding his behavior outside the ritual space.

3 Understanding Ritual

This discussion of archaic rituals from early China may seem – at best – interesting as an exercise in historical archaeology. But I would like to argue that there is much here that is good to think concerning norms in general.

So how should we interpret these rituals?

One way – and perhaps the most common way in contemporary theory – would be to assume that ritual serves to inculcate a worldview that the practitioner is being called upon to accept. From such a perspective, the goal of the analyst would be to explore the views that are embedded in the ritual and try to work out the larger worldview in which such ideas would make sense.

In the case at hand, the goal of the ritual would be read as trying to inculcate a proper reverence for the deceased as an ancestor. This would then be analyzed in terms of a larger worldview in which humans would be taught to see themselves as following the dictates of ancestral powers.

This is, in fact, the most common way these early Chinese ancestral rituals have been interpreted by modern scholars. But, as I have argued elsewhere, it is a misreading.[10] The *Book of Rites* presents ancestors as a product of ritual, and they

[9] *Liji*, "Za ji," ICS, 107/20.12/6.

[10] Puett, "The Offering of Food and the Creation of Order: The Practice of Sacrifice in Early China," in Roel Sterckx, ed., *Of Tripod and Palate: Food, Politics, and Religion in Traditional China* (New York: Palgrave MacMillan, 2005), pp. 75–95.

are revered as such only within the ritual space. Outside the ritual space, they are regarded as ghosts, and potentially very dangerous ghosts.

In other words, one should not try to make sense of rituals by working out the larger worldview that a ritual is supposedly trying to inculcate. On the contrary, rituals often work because they are *counter-intuitive* to our usual ways of thinking, acting, and responding. And it is precisely the tension between rituals and our lived reality that render them effective.

4 The "As If" World of Ritual

For our concerns, one of the more significant statements in the early Chinese tradition concerning ritual appears in the *Analects* of Confucius:

> 'Sacrifice as if present,' means: sacrifice to the spirits as if the spirits were present. The master said, ''If I do not participate in the sacrifice, it is as if I did not sacrifice.''[11]

The "as if" world of the ritual involves the presence of both the human practitioner and the spirits. Within this "as if" world, the practitioner and the spirits are in close proximity, and the practitioner works to forge a proper relationship between the descendant and ancestor. But whether the spirits are actually present or not is irrelevant. The ritual rather serves as a space within which one acts "as if" they are present.

Moreover, as is clear from other statements, this ritual space is clearly demarcated from the non-ritual world. It may not matter if the spirits are present or not in the ritual space. But one actually does *not* want them to be present, or at least not closely so, outside the ritual space. As the *Analects* states elsewhere:

> Fan Chi asked about knowledge. The master said, "To work on behalf of what is proper for the people, to be reverent to the ghosts and spirits and yet keep them at a distance, this can be called knowledge."[12]

One wants the presence (whether actual or not) of the spirits in the "as if" world of the ritual space; outside it, one wants the spirits kept at a distance.

Within the ritual space, one develops the proper dispositions one should have in a relationship between an ancestor and a descendant. This helps to re-inculcate into oneself the sensibility of being a proper descendant. Outside the ritual space, one is guided by separate dispositions.

The "as if" world of ritual is clearly marked off as such, demarcated from the world of lived reality outside the ritual space. It is not a world we could live in. But, perhaps more importantly, we would not *want* to live in the "as if" world, even if this were possible. The "as if" world is not a simple repository of norms that could guide our behavior in the world of our lived reality. The power of the "as if" on the

[11] *Lunyu*, 3/12.

[12] *Lunyu*, 6/22.

contrary resides in the productive tension that it generates with the world of our lived experience.

Several colleagues and I have argued that early Chinese understandings of ritual should be taken seriously from a theoretical perspective as well. In these understandings, ritual serves to create an "as if" space – what we call a subjunctive space – in which dispositions can be trained.[13] What is of particular significance is that, in such understandings, normative values are located not in the "as if" space itself but rather in the disjunction between the "as if" space and our experiences outside of the ritual space:

> These arguments imply that ritual always operates in a world that is fragmented and fractured. Moreover, the subjunctive world created by ritual is always doomed ultimately to fail – the ordered world of flawless repetition can never fully replace the broken world of experience. This is why the tension between the two is inherent and, ultimately, unbridgeable. Indeed, this tension is the driving force behind the performance of ritual: the endless work of ritual is necessary precisely because the ordered world of ritual is inevitably only temporary…. If the world is always fractured, and if ritual always operates in tension with such a world, then we need to think of ritual in terms of such an endlessly doomed dynamic. Ritual should be seen as operating in, to again quote Robert Orsi, "the register of the tragic." Although the claims of ritual may be of an ordered, flawless system, the workings of ritual are always in the realm of the limited and the ultimately doomed.[14]

5 Where are the Norms?

Let us think through the larger implications of the position under discussion here. We often tend to think of norms in the form of rules to guide our behavior. And having such norms embedded in rituals would appear to be extremely limiting. To begin with, we frequently assume that rituals provide a type of normative order to which we should try to approximate our behavior. If this is really the way rituals function, then they would simply be telling us what to do. Moreover, as mentioned above, having such rules embedded in rituals would seem to limit the potentials for altering those norms: if the rules embedded in rituals are based upon traditional values, and if rituals serve to inculcate those traditional values within ourselves, then they would be extremely difficult to alter.

But the vision being developed by the ritual theorists from classical China opens up a different way to think about rituals and, by implication, a different way to think

[13] Adam Seligman, Robert Weller, Michael Puett, and Bennett Simon, *Ritual and its Consequences: An Essay on the Limits of Sincerity* (Oxford: Oxford University Press, 2008). For a superb analysis of "as if" from a larger philosophical perspective, see Hans Vaihinger, *The Philosophy of 'As if': A System of the Theoretical, Practical and Religious Fictions of Mankind*, translated by C. K. Ogden, second edition (New York: Harcourt, Brace and Company, 1935).

[14] Adam Seligman, Robert Weller, Michael Puett, and Bennett Simon, *Ritual and its Consequences: An Essay on the Limits of Sincerity* (Oxford: Oxford University Press, 2008), p. 30.

about norms. Rituals do not, in these theories, teach us what to do; they rather, to refer back to the statement quoted above, operate in the realm of the tragic: they operate in the very disjunction between the "as if" world of ritual and our lived reality.[15]

The ritual of ancestor worship is not a norm in the sense of something that we blindly follow: the ritual is not inculcating a belief in ancestors. Instead, it works because the relationships that are being built in the ritual space are in tension with the relationships that exist outside the ritual space.

6 What are the Norms?

These ritual theories assume a particular vision of the self. As humans, we tend to fall into sets of patterned responses to the world. The goal of rituals – of "as if" spaces – is to break us out of these patterns. However, it is not the case that these as-if spaces are therefore the locus of normativity – in other words, that ritual is functioning to socialize us into the norms embodied in the ritual. The rituals are not telling us what to do; our patterned set of responses is rather guiding our everyday lives. The point of ritual is to break us out of these patterns.

According to the *Analects*, the goal of this work – of working through the disjunctions created between ritual and the world of our lived experience – is to create humaneness (*ren* 仁). Humaneness is notoriously difficult to explicate precisely, because it is a sensibility – a sensibility of how to act in ways that will help those around one. In one passage of the *Analects*, Confucius is quoted as defining humaneness, vaguely but tellingly, as "Caring for others."[16] And, of course, what this means in any particular situation will vary dramatically. But the reason we need to work through rituals to attain this sensibility is because, according to this way of thinking, we on the contrary fall into patterns in our responses to the world, and thus cease to be able to sense how to act humanely. The work of ritual – of breaking these patterns by entering "as if" spaces – is thus in a sense a training exercise in which we are training ourselves to respond to situations better.

In the case of the mourning rituals, the goal is not to inculcate a particular view of ancestors. The goal of the rituals is to break us from our tendency to fall into dangerous patterns at the death of a loved one and to help us channel these dispositions more productively. Out of the disjunction between these two will hopefully come a more refined way of responding to the world.

Ritual norms, in other words, are not rules that we simply follow. Rituals rather work because they create sets of patterned responses that stand in tension with the patterned responses we have in our lived reality. It is, in short, the disconnect between rituals and lived reality that makes rituals so effective.

[15] Michael Puett, "Ritual Disjunctions: Ghosts, Philosophy, and Anthropology," in Das et al., eds., op. cit., pp. 218–23. See also "Critical Approaches to Religion in China," *Critical Research on Religion* 1 (2013): 95–101.

[16] Lunyu, 12.22.

It is not the case, then, that the norms are to be found in the rituals, with the rituals forming a kind of normative set of rules that should guide our behavior. Rituals on the contrary operate to force us out of our patterned responses to the world. But the patterns that rituals offer are hardly ones that we would always be following either. The guiding norm is rather a sensibility of responding to situations well – a sensibility we attain by working through the disjunctions between our common lived behavior and the as-if worlds of rituals. Rituals are not the repository of norms but rather the ever-altering means to achieve norms.

7 Conclusion

The understandings of ritual that we have been exploring open up a possible way of thinking about norms. According to the ritual theorists under discussion here, norms are located not in a series of rules that should be followed. Norms are rather to be found in the sensibilities that are developed over time, and the way that such norms are developed is through the tension between rituals and our lived reality. We are constantly creating pockets of "as if" realities, and the disjunction between these pockets and the patterns that we fall into in our lived reality is the basis for us to transform ourselves. But what we are seeking is not to become more like the person we are in these as-if spaces. The goal is rather to learn to respond to situations well – an ability we gain through the endless work of training ourselves through ritual activity.

From this perspective, rituals are best thought of as patterned set of responses operating in tension with the patterned set of responses that usually govern our lives. And norms should rather be thought of as more generalized (and often quite ill-defined) concepts such as caring. Rituals are then, in a sense, a way of training ourselves to break from those patterns that usually prevent us from being caring toward others.

Taking these theories from classical China seriously allows us to view rituals, and, ultimately, norms in a different way. If rituals are thought of not as a repository of norms into which we are being inculcated but rather as a series of "as if" worlds that stand in tension with our lived reality, then they become the means through which we train ourselves to become more able to care for those around us.

frontiers in
PSYCHOLOGY

Normativity, interpretation, and Bayesian models

PERSPECTIVE ARTICLE
published: 15 May 2014
doi: 10.3389/fpsyg.2014.00332

Normativity, interpretation, and Bayesian models

Mike Oaksford*

Department of Psychological Sciences, Birkbeck College, University of London, London, UK

***Correspondence:**
Mike Oaksford, Department of Psychological Sciences, Birkbeck College, University of London, Malet Street, London WC1E 7HX, UK
e-mail: mike.oaksford@bbk.ac.uk

Edited by:
David E. Over, Durham University, UK

Reviewed by:
Ulrike Hahn, University of London, UK
Vincenzo Crupi, University of Turin, Italy

It has been suggested that evaluative normativity should be expunged from the psychology of reasoning. A broadly Davidsonian response to these arguments is presented. It is suggested that two distinctions, between different types of rationality, are more permeable than this argument requires and that the fundamental objection is to selecting theories that make the most rational sense of the data. It is argued that this is inevitable consequence of radical interpretation where understanding others requires assuming they share our own norms of reasoning. This requires evaluative normativity and it is shown that when asked to evaluate others' arguments participants conform to rational Bayesian norms. It is suggested that logic and probability are not in competition and that the variety of norms is more limited than the arguments against evaluative normativity suppose. Moreover, the universality of belief ascription suggests that many of our norms are universal and hence evaluative. It is concluded that the union of evaluative normativity and descriptive psychology implicit in Davidson and apparent in the psychology of reasoning is a good thing.

Keywords: psychology of reasoning, Bayesian models, Bayesian argumentation, radical interpretation, Donald Davidson, evaluative normativity

Elqayam and Evans (2011) have argued against *evaluative* normativity having any role in psychological theories of reasoning. They contrast evaluative normativity with *directive* normativity. They argue that directive normativity is conditional and perfectly

consistent with programs in cognitive science like rational analysis (Anderson, 1990; Oaksford and Chater, 1998, 2007). Consequently, they have no problem with formulations like, *if you want to be well adapted to your environment then you should act in a Bayes optimum fashion in classification, decision and prediction*. However, what we can't apparently assert is the unconditional *you should act in a Bayes optimum fashion in classification, decision, and prediction*. This is an evaluative claim suggesting in some absolute sense that this is the right way to behave. In particular, they observe that if there were an alternative normative theory of what constitutes being well adapted to your environment, citing empirical evidence to distinguish between these two normative theories would commit the *is-ought* fallacy. Consequently, evaluative normativity should be expunged from psychological theorizing about reasoning.

In this paper, I pursue a broadly Davidsonian (Davidson, 2004) response to Elqayam and Evans' (2011). In the first section, *Types of Rationality*, I set up the argument by observing that two distinctions they make, between *instrumental* and *normative* rationality and between *directive* and *evaluative* rationality, are far more permeable than they require. I conclude that Elqayam and Evans (2011) primary objection is to the suggestion that we should pick the theory that makes the most rational sense of our data. In the second section, *Interpretation, Argumentation, and Rationality*, I argue that this is inevitable consequence of Davidson's account of radical interpretation. On Davidson's view, rationality is a social construct where to interpret others' statements requires that we adopt a principle of charity, i.e., they share the same norms as ourselves. Davidson's account suggests attributing people with intentional states like beliefs requires evaluative normativity. I then show that in the social context of argumentation, a third person argument evaluation methodology yields close conformity to rational Bayesian norms. Participants are quite capable of evaluating others arguments. I conclude that this ubiquitous human behavior is something that psychology must explain. In the final section, *How Many Rational Norms Are There?* I argue that logic and probability theory are not really competing norms, the important psychological question is whether beliefs are binary or graded. Moreover, following Davidson, I question Elqayam and Evans (2011) grounds for normative relativism. In conclusion, I suggest that while there are many outstanding problems and exceptions, the continuing union of evaluative normativity and descriptive psychology apparent in the psychology of reasoning is a good thing.

TYPES OF RATIONALITY

Stanovich (2011) argued that Elqayam and Evans (2011) drive a wedge between Bayesian probability theory, which they regard as an account of normative rationality, and *instrumental* rationality. Instrumental or practical rationality, which Elqayam and

Evans (2011) endorse, provides a suitable means for achieving one's goals regardless of the nature of those goals. However, as Stanovich (2011) observes, this is a difficult wedge to drive home given that the standard justification for the laws of subjective probability are given by the Dutch book theorem (Vineberg, 2011). For each of the laws of probability theory, this theorem establishes that violating them would leave an agent open to making bets they cannot win. The *converse* Dutch book theorem then establishes that these laws are instrumentally rational because conforming to them prevents taking self-defeating actions. This instrumentally rational justification can then be provided with a directively rational formulation: *if an agent wishes to avoid making bets they cannot win, then they should conform to the laws of probability theory*. This conditional formulation just restates the *converse* Dutch book theorem. So this formulation involves making conformity to the normative theory conditional on that normative theory's rational justification. The justification for probability theory is instrumental [other epistemic justifications, based on maximizing accuracy, are equally instrumental (Joyce, 1998)]. So, in the case of probability theory there is simply no wedge to be driven between instrumental and normative rationality[1].

This formulation also raises the question of how universal are the goals stated in the antecedent? In a conditional formulation the more universal an antecedent the less it needs to be stated. So, for example we would normally say *ripe apples fall*. We do not feel compelled to formulate this as *if gravity is in force ripe apples fall*. One could even use an appropriate modal, *ripe apples ought to fall*. Certainly one might be inclined to query whether this is a real or a good apple if it did not fall, which is perilously close to an evaluative judgment. Similarly, the more universal we regard the wish to avoid making bets one is bound to lose, the more inclined we would be to drop the conditional formulation and evaluate anyone not conforming to the rules of probability as irrational just as we may be inclined to evaluate the apple as inedible. If we encountered someone willing to make bets they were bound to lose, they would probably be institutionalized for their own safety. As with instrumental and normative rationality, the barrier between directive and evaluative rationality seems permeable. Moreover, the fundamental issue is of universality versus relativity. The theory is normatively rational if its justification is considered universal.

The inference to which Elqayam and Evans (2011) seem to take exception is the claim that as theoreticians we *should* accept the theory that makes the most rational sense of the participants' behavior (Oaksford and Chater, 1996, 2007). As long as we are comparing the rules of normative theories, this will mean that the one that best describes participants' behavior is the one that makes most rational sense of it. This thesis derives from the fact that in interpreting empirical data, i.e., our participants' behavior, we are in exactly the same position as the *radical* interpreter in Davidson's (1984, 2004) theory of ascribing intentional content. The difference is that as reasoning researchers we may have more than one normative theory in mind, whereas in radical

interpretation one imputes one's own norms to one's interlocutor. However, the general principle remains the same: we are trying to make the best sense of what we have been told.

INTERPRETATION, ARGUMENTATION, AND RATIONALITY

Davidson's model of radical interpretation is an idealized account of how a cognitive agent can interpret another agent's behavior and utterances to infer their beliefs and desires (Rescorla, 2013). The model is based on Bayesian decision theory, in which beliefs are graded and related to subjective probabilities and people's desires are represented as utilities. Savage's (1954) axioms show that when a person's preferences meet certain requirements there are probabilities and utilities that guarantee that their preferences maximize expected utility. Consequently, an agent's beliefs and desires can be inferred from their overt preferences. An important wrinkle is that the propositional content of beliefs are not pre-specified but are also inferred from an interlocutor's preferences for the truth of sentences. Central to this account is the thesis that to ascribe another person with the appropriate beliefs and desires means we must assume they conform to our own standards of rationality. This is the principle of charity. As Davidson (2005; p. 319, cited in, Rescorla, 2013) puts it: "Charity is a matter of finding enough rationality in those we would understand to make sense of what they say and do, for unless we succeed in this, we cannot identify the contents of their words and thoughts." Rationality is constitutive of having intentional states.

This is an idealized model but the central idea that we must attribute to others similar rational norms to ourselves in order to interpret them is intended as a more general claim about interpretation in the real world that involves attributing others with propositional attitudes like beliefs and desires. On Davidson's view describing somebody's behavior in terms of beliefs and desires is inseparable from normative evaluation.

Davidson's (2004) emphasis on interpretive communicative processes proposes a particular research methodology which has been pursued recently in the context of human argumentation (Oaksford and Hahn, 2004, 2013; Hahn and Oaksford, 2007). Argumentation is a social phenomenon in which one or more people attempt to persuade another person or group of a particular, often controversial, position. It is a commonplace of argumentation theory that arguing is pointless unless there is broad agreement between the protagonists on what could count as a reasonable argument (Perelman and Olbrechts-Tyteca, 1969; Woods et al., 2004). Without this point of departure there is no point in engaging in an argumentative exchange. At least initially, we must apply the principle of charity[2]. Recently it has been argued that reasoning usually has an argumentative goal (Hahn and Oaksford, 2007; Mercier and Sperber, 2011). Consequently, it is in social argumentative contexts where people's rational

norms of reasoning would be expected to be most in evidence. It is a critical ability to be able to evaluate the arguments put forward by others to persuade you or your friends of particular positions.

Recent research in this area has adopted a third person argument evaluation methodology (Oaksford and Hahn, 2004, 2013; Hahn and Oaksford, 2007; Harris et al., 2012). Participants are explicitly asked to assess the degree to which one interlocutor, A, *should* be convinced by an argument put forward by another interlocutor, B. So, participants are explicitly asked for an evaluative judgment. They are also provided with information about A's prior degree of belief in the conclusion. Hahn and Oaksford (2006, 2007), Oaksford and Hahn (2004, 2013) have provided normative Bayesian analyses of a variety of different forms of argumentation which make clear predictions for participants' judgments. In this context, a normative Bayesian account provides excellent fits to the data. Moreover, this is true even when there are no parameters free to vary (Harris et al., 2012) because participants have been asked for their judgments of the relevant likelihoods from which predictions for their posteriors can be directly computed (see also, Fernbach and Erb, 2013). These results demonstrate that when participants are asked for an evaluative judgment of other peoples' arguments they reveal behavior that is closely in accordance with the appropriate normative theory. This is not only because they have been asked directly to make an evaluative judgment. They are also explicitly provided with A's prior degree of belief, which absolves them from the dilemma of considering whether *they* would believe the conclusion prior to hearing the argument. They are simply told that, for whatever reason, A believes it to a certain degree. In first person paradigms, participants are asked to assume or suppose that *they* believe the premises to be true or to a certain degree, when of course they may believe no such thing.

In summary, the psychology of reasoning will have to deal with evaluative normativity because much human behavior involves the explicit evaluation of others' arguments, especially in politics, and in the law. Moreover, participants in experiments on argumentation make these evaluations naturally and their performance reveals direct sensitivity to appropriate rational norms.

HOW MANY RATIONAL NORMS ARE THERE?

I conclude this paper by addressing two critical issues underlying Elqayam and Evans (2011) criticisms of evaluative normativity, (i) deciding between normative theories and (ii) the conviction that constructs like the principle of charity collapse into relativism. On Davidson's (2004) ideal model there are no alternative normative frameworks. Basic logic, probability theory, and decision theory [see, Chater and Oaksford (2012) on the role of these theories in cognitive science] are fundamental rational norms and he broaches no other possibilities. This raises the question, of how

many rational norms are there actually to choose between? A prima facie argument can be made that that there are not as many as one might think. Elqayam and Evans (2011) suggest that the new Bayesian paradigm is an alternative norm account. I argue that since probability theory presupposes standard logic they are not really in competition. A derived theorem of the Kolmogorov axioms is *logical consequence*, i.e., *if X logically entails Y, then* $Pr(Y) \geq Pr(X)$, which "ensures that probabilistic reasoning respects deductive logic" (Joyce, 2004, p. 135). The question is not whether one norm supplants another but whether beliefs are graded. Once we opt for graded beliefs, then we need to know how they are updated in inference when new information comes in. This can be achieved by Bayesian conditionalization rather than modus ponens (Oaksford, in press; Oaksford and Chater, 2007, 2013), although this is not necessary because probabilistic premises will *deductively* entail a probability interval for the conclusions of an argument (Pfeifer and Kleiter, 2010). Consequently, I suggest that the move to Bayesian probability is not a move to an alternative norm rather than a move to a finer grained analysis of beliefs which is not just binary true or false.

Thus, when comparing logic and probability, we are not choosing between competing norms. Davidson would argue, and common sense seems to dictate, that if the more nuanced view provides a rational understanding of more of the data it is the preferred theory. When the issue of competing norms is taken out of the equation this is simply the question of which theory provides the best description of the data. What happens if there are genuinely competing normative theories that are equally descriptively adequate?

For example, in decision theory an explicit competitor to classical Bayesian probability theory has been provided by quantum probability (Pothos and Busemeyer, 2013). This would appear to be much closer to the competing norms case that Elqayam and Evans (2011) envisage. Quantum probability stands to quantum logic – in which the law of the excluded middle is not valid – as Bayesian probability stands to standard logic (Oaksford, 2013). Moreover, across a variety of tasks, Pothos and Busemeyer (2013) argue that quantum probability is more descriptively adequate than Bayesian probability theory. Recall that the formulation for directive normativity is conditional, with the relevant justification for the normative theory in the antecedent. For Bayesian probability theory we have, *if an agent wishes to avoid making bets they cannot win, then they should conform to the laws of probability theory*. For quantum probability, however, there does not appear to be a relevant justificatory antecedent. There would appear to be no Dutch book theorem showing that failure to conform to the laws of quantum probability would lead anyone to make bets they could not win[3]. Moreover, conformity to the laws of quantum probability may well lead to a Dutch book being made against you. For example, it has been shown that committing the conjunction fallacy (Tversky and Kahneman, 1983) can allow a Dutch book to be made against you (Gilio and Over, 2012; Hahn 2014) and quantum probability apparently *predicts* the

conjunction fallacy (Pothos and Busemeyer, 2013). Consequently, however descriptively adequate with respect to the data quantum probability appears to be, it cannot explain how behavior succeeds in the real macroscopic world which we inhabit. Even if we can make sense of laying bets on the outcomes of quantum events, there would still need to be an independent argument that there are similar events about which we could gamble at the macroscopic level (Oaksford, 2013).

Elqayam and Evans (2011) argue against the principle of charity solely on the observation that norms are relative to particular cultural and historical contexts. However, they do not discuss Davidson's view of rationality as a constitutive norm (Rescorla, 2013). On Davidson's view conformity to these norms is constitutive of having intentional states and is not relative to any particular cultural or historical context. As there are no human beings to whom we would not attribute beliefs this suggests that our norms are also universal. The Dutch book theorems certainly have this character. Gambling is a universal human activity, engaged in by all cultures and in all historical contexts. Moreover, it seems inconceivable that anyone would fail to accede to the rationale for the Dutch book theorems, what normal human being would wish to make bets they are bound to lose? In the first section, I argued that the permeability between directive and evaluative rationality depends on the universality of the justification for a normative system. So we have good grounds to view probability theory as a universal evaluative norm.

CONCLUSION

In conclusion, in the psychology of reasoning, interpreting experimental results, just as in interpreting another's utterances, requires making the best rational sense of the observed behavior. People evaluate each other's arguments in politics and in the law and in appropriate argumentative contexts their judgments conform to the rational norms of probability theory. The current Bayesian turn in the psychology of reasoning addresses the question of whether beliefs are graded and is not an alternative norm to standard logic. From Davidson's perspective, the universal attribution of beliefs to others has the corollary that our rational norms are likely to be similarly universal. Elqayam and Evans (2011) provide no grounds to question this perspective. However, there are many exceptions, data that does not conform to these norms (e.g., Tversky and Kahneman, 1983; but see, Crupi et al., 2008), cases of irrationality due to illness or injury, cases where sacred values are opposed to utility maximization (Atran and Axelrod, 2008), and other paradoxes of maximizing expected utility (Burns and Wieth, 2004; but see Turner and Quilter, 2014). However, there are responses to these exceptions as some of the citations indicate. In sum, the union of evaluative normativity and descriptive psychology, implicit in Davidson (Rescorla, 2013), is continuing to yield important results and this should be regarded as a good thing.

Footnotes

[1]We note also that the justification for selecting data in accordance with Oaksford and Chater's (1994) information gain model is again instrumental. So following its dictates will mean that this strategy minimizes the length of the sequential sample needed for the posteriors to converge on the true hypothesis (Fedorov, 1972). This is an instrumental justification: *if* people want to get to the truth in the most economical way they will select data in accordance with the theory.

[2]After an initial exchange, we may discover that we are not in a critical discussion, i.e., a rational exchange of arguments intended to persuade, but rather are in a quarrel, where rationality goes out the window.

[3]Although in physics, there are arguments that a Bayesian approach, i.e., probability as a measure of ignorance, might make sense of quantum probability as a theory of rational betting in quantum gambles (Pitowsky, 2003). One then has to ask whether there is any analog of a quantum gamble at the macroscopic level that any human being would be concerned to win.

REFERENCES

Anderson, J. R. (1990). *The Adaptive Character of Thought*. Mahwah, NJ: Erlbaum.

Atran, S., and Axelrod, R. (2008). Reframing sacred values.*Negotiation J.* 24, 221–246. doi: 10.1111/j.1571-9979.2008.00182.x

Burns, B. D., and Wieth, M. (2004). The collider principle in causal reasoning: why the monty hall dilemma is so hard *J. Exp. Psychol. Gen.* 133, 434–449. doi: 10.1037/0096-3445.133.3.434

Chater, N., and Oaksford, M. (2012). "Normative systems: logic, probability, and rational choice," in*The Oxford Handbook of Thinking and Reasoning*, eds K. Holyoak and R. Morrison (Oxford: Oxford University Press), 11-21.

Crupi, V., Fitelson, B., and Tentori, K. (2008). Probability, confirmation, and the conjunction fallacy. *Think. Reason.* 14, 182–199. doi: 10.1080/13546780701643406

Davidson, D. (1984). "On the very idea of a conceptual scheme," in*Inquiries into Truth and Interpretation,* ed. D. Davidson (Oxford: Oxford University Press), 183–198.

Davidson, D. (2004). *Problems of Rationality*. Oxford: Clarendon Press. doi: 10.1093/0198237545.001.0001

Davidson, D. (2005). *Truth, Language, and History*. Oxford: Clarendon Press. doi: 10.1093/019823757X.001.0001

Elqayam, S., and Evans, J. T. (2011). Subtracting 'ought' from 'is': descriptivism versus normativism in the study of human thinking. *Behav. Brain Sci.* 34, 233–248. doi: 10.1017/S0140525X1100001X

Fedorov, V. (1972). *Theory of Optimal Experiments*. London: Academic Press.

Fernbach, P. M., and Erb, C. D. (2013). A quantitative causal model theory of conditional reasoning.*J. Exp. Psychol. Learn. Mem. Cogn.* 39, 1327–1343. doi: 10.1037/a0031851

Gilio, A., and Over, D. (2012). The psychology of inferring conditionals from disjunctions: a probabilistic study.*J. Math. Psychol.* 56, 118–131. doi: 10.1016/j.jmp.2012.02.006

Hahn, U. (2014). The Bayesian boom: good thing or bad? *Front. Psychol.* (in press).

Hahn, U., and Oaksford, M. (2006). A Bayesian approach to informal argument fallacies.*Synthese* 152, 207–236. doi: 10.1007/s11229-005-5233-2

Hahn, U., and Oaksford, M. (2007). The rationality of informal argumentation: a Bayesian approach to reasoning fallacies. *Psychol. Rev.* 114, 704–732. doi: 10.1037/0033-295X.114.3.704

Harris, A. L., Hsu, A. S., and Madsen, J. K. (2012). Because hitler did it! Quantitative tests of Bayesian argumentation using ad hominem. *Think. Reason.* 18, 311–343. doi: 10.1080/13546783.2012.670753

Joyce, J. M. (1998). A nonpragmatic vindication of probabilism. *Philos. Sci.* 65, 575–603. doi: 10.1086/392661

Joyce, J. M. (2004). "Bayesianism," in*The Oxford Handbook of Rationality,* eds A. R. Miele and P. Rawling (Oxford: Oxford University Press), 132–155. doi: 10.1093/0195145399.003.0008

Mercier, H., and Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *Behav. Brain Sci.* 34, 57–74. doi: 10.1017/S0140525X10000968

Oaksford, M. (2013). Quantum probability, intuition, and human rationality. *Behav. Brain Sci.* 36, 303. doi: 10.1017/S0140525X12003081

Oaksford, M. (in press). "Knowing enough to achieve your goals: Bayesian models and practical and theoretica rationality in conscious and unconscious inference," in *Human Rationality: Thinking Thanks to Constraints*, eds L. Macchi, L. M. Bagassi, and R. Viale (Cambridge, MA: MIT Press).

Oaksford, M., and Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychol. Rev.* 101, 608–631. doi: 10.1037/0033-295X.101.4.608

Oaksford, M., and Chater, N. (1996). Rational explanation of the selection task. *Psychol. Rev.* 103, 381–391. doi: 10.1037/0033-295X.103.2.381

Oaksford, M., and Chater, N. (eds). (1998). *Rational Models of Cognition*. Oxford: Oxford University Press.

Oaksford, M., and Chater, N. (2007). *Bayesian Rationality: The Probabilistic Approach to Human Reasoning* Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780198524496.001.0001

Oaksford, M., and Chater, N. (2013). Dynamic inference and everyday conditional reasoning in the new paradigm *Think. Reason.* 19, 346–379. doi: 10.1080/13546783.2013.808163

Oaksford, M., and Hahn, U. (2004). A bayesian analysis of the argument from ignorance. *Can. J. Exp. Psychol.* 58, 75–85. doi: 10.1037/h0085798

Oaksford, M., and Hahn, U. (2013). "Why are we convinced by the ad hominem argument?: Bayesian source reliability and pragma-dialectical discussion rules," in *Bayesian Argumentation,* ed. F. Zenker (Dordrecht: Springer), 39–59.

Perelman, C., and Olbrechts-Tyteca, L. (1969). *The New Rhetoric: A Treatise on Argumentation*. Notre Dame, IN University of Notre Dame Press.

Pfeifer, N., and Kleiter, G. (2010). "Mental probability logic," in *Cognition and Conditionals: Probability and Logic in Human Thinking,* eds M. Oaksford and N. Chater (Oxford: Oxford University Press), 153–173. do 10.1093/acprof:oso/9780199233298.003.0009

Pitowsky, I. (2003). Betting on the outcomes of measurements: a Bayesian theory of quantum probability. *Stud. Hist. Philos. Mod. Phys.* 34, 395–414. doi: 10.1016/S1355-2198(03)00035-2

Pothos, E. M., and Busemeyer, J. R. (2013). Can quantum probability provide a new direction for cognitive modeling *Behav. Brain Sci.* 36, 255–274. doi: 10.1017/S0140525X12001525

Rescorla, M. (2013). "Rationality as a constitutive ideal," in *A Companion to Davidson,* eds E. Lepore and K. Ludwig (Oxford: Wiley-Blackwell), 472–488. doi: 10.1002/9781118328408.ch27

Savage, L. (1954). *The Foundations of Statistics*, 2nd Edn. New York: Dover.

Stanovich, K. E. (2011). Normative models in psychology are here to stay. *Behav. Brain Sci.* 34, 268–269. doi: 10.1017/S0140525X11000161

Turner, R., and Quilter, T. (2014). *The Two Envelopes Problem* (*Gatsby Centre for Computational Neuroscience*). Available at: http://www.gatsby.ucl.ac.uk/~turner/Notes/TwoEnvelopes/2envlps.pdf [accessed January 3, 2014].

Tversky, A., and Kahneman, D. (1983). "Extensional vs. intuitive reasoning: the conjunction fallacy in probability judgement," in *Heuristics and Biases: The Psychology of Intuitive Judgement,* eds T. Gilovich, D. Griffin, and D. Kahneman (New York: Cambridge University Press), 19–48.

Vineberg, S. (2011). "Dutch book arguments," in *The Stanford Encyclopedia of Philosophy,* Summer 2011 Edn, ed. E. N. Zalta. Available at: http://plato.stanford.edu/archives/sum2011/entries/dutch-book/ [accessed January 4, 2014].

Woods, J., Irvine, A., and Walton, D. N. (2004). *Argument: Critical thinking, Logic and the Fallacies* (Revised Edition). Toronto, ON: Prentice Hall.

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. The Associate Editor declares that while the author Mike Oaksford as well as the reviewer Ulrike Hahn are currently affiliated with the same

department (Department of Psychological Sciences, Birkbeck College, University o London), there has been no conflict of interest during the review and handling of this manuscript.

Received: 05 February 2014; accepted: 31 March 2014; published online: 15 May 2014.

Citation: Oaksford M (2014) Normativity, interpretation, and Bayesian models. Front. Psychol. ***5****:332. doi: 10.3389/fpsyg.2014.00332*

This article was submitted to Frontiers in Cognitive Science, a specialty of Frontiers in Psychology.

Normative Analysis in the Communications Field: Why We Should Distinguish Communicative Means and Ends of Justice

Author(s): Max Hänska

Source: *Journal of Information Policy*, 2019, Vol. 9 (2019), pp. 56-78

Published by: Penn State University Press

Stable URL: https://www.jstor.org/stable/10.5325/jinfopoli.9.2019.0056

NORMATIVE ANALYSIS IN THE COMMUNICATIONS FIELD

Why We Should Distinguish Communicative Means and Ends of Justice

Max Hänska

ABSTRACT

In the social sciences, we often face normative questions, not least because many areas of inquiry intersect with public policy. Understanding and explaining media and communications is one task, deciding how communication systems *should* be organized quite another, but normative analysis receives scant attention. This article explores normative analysis: what is involved in answering questions about justice and communication, about how sociopolitical and indeed communicative arrangements ought to be organized.

Keywords: communication, technology, justice, normative analysis

Habermas[1] famously lamented that the rise of publicity and entertainment was refeudalizing the public sphere. The work of Hall[2] and many others has shown us how representations can significantly affect the social standing, opportunities, and even rights of minorities. More recently, a host of issues around privacy, ownership, and control of private data have gained widespread attention. Consider the myriad ways in which large platforms with deep pools of data can affect how people are represented, made visible, treated, and gain voice[3]; how injustices materialize in networked publics[4]; or how spirals of silence, trolls, and botnets can drown out legitimate voices, and become a mode of censorship.[5] The list, of course, can go on.

Max Hänska: Conflict and Civil Society Research Unit, London School of Economics

1. Habermas.
2. Hall and Gates.
3. Taylor.
4. Harvey and Leurs.
5. Noelle-Neumann; Tufekci, *Twitter and Tear Gas*.

JOURNAL OF INFORMATION POLICY, Volume 9, 2019

The point is, scholarly and public discourse are suffused with appraisals of what is good and bad about communications, its technologies, and other communicative and informational resources. Our discourse is laced with reference to media and democracy, data and justice, communication and resistance, voice and representations of issues and minorities, media monopolies and power, the myriad ways in which social media are said to empower individuals, and warp democratic politics.[6] These appraisals tread the peculiar terrain of normative thought. In a broader sense, they raise questions about what, if anything, justice requires of communication.

This essay does not answer that question, at least not directly. Rather, it will map the place of normative analysis in the communication field. By normative analysis I mean the ways we construct and justify normative arguments. While there are extensive discussions about methods of empirical social research, the question what makes a normative claim plausible, robust, and worth endorsing receives slim to no attention. Even in political philosophy, normative analysis receives very little systematic explication. In the communications field claims about what is good and desirable, or bad and pernicious are usually asserted, such that the reader should already accept their rightness. Perhaps the distinction between right and wrong seems too obvious to us, to merit systematic attention. But I want to suggest, that while critical appraisals of bad things we are against may come easily to us, it is much harder to say with precision what we are for. That is why this article aims to set out some basic parameters of normative analysis. What is involved in making assessments about the justness of communicative processes, practices, technologies, institutions, and other informational and communicative resources?

The focus will primarily be on communications, by which I mean processes and practices of producing, expressing, disseminating, accessing, and receiving information and cultural artifacts. As communication practices and processes are often inseparable from technologies, media institutions, and communicative and informational resources, the discussion will touch upon these too. I have in mind questions about people's ability to communicate, to express their views, to be heard, and to access information—questions about information quality, control over information, and representations.

6. Althaus; Bucy and D'Angelo; Gerhards and Neidhardt.

Normativity itself requires a brief definition. Tufekci recently wrote about the problem of data ownership and privacy that "we should all be thinking of ways to reintroduce competition into the digital economy. Imagine, for example, requiring that any personal data you consent to share be offered back to you in an 'interoperable' format, so that you could choose to work with companies you thought would provide you better service."[7] In the United Kingdom, a group of academics called on the press to "Stop Jeremy Corbyn's trial by media."[8] What these quotes have in common, is that they express normative ideas. For Scanlon, normative questions are about the things that people have reasons to do. Thomson[9] distinguishes between normative propositions that are evaluatives and directives. Tufekci's statement is a directive: this is what we *ought* to be doing. The second statement is both evaluative and directive: what the media doing is bad. They should stop it. Both imply that there are good and important reasons for us to act in a different manner—even if the underlying reasons are not fully explicated.

At its simplest, normative statements (and theories) can be contrasted with empirical statements (and theories). While the latter describes the world as it *is* believed to be, the former tells us how it *ought* to be. The concept of "direction of fit" can elucidate the distinction.[10] Let us define theory as mental models (representations of the world, the way we think about or imagine something). Empirical theories have a mind-to-fit-world direction of fit, because they express a belief about how the world *is* believed to be. They aim to explain, or understand the world as it actually *is*. That is, we want our theories, our mental models, to represent (fit) the world as accurately as possible. Normative theories have a world-to-fit-mind direction of fit, because they express a belief about how the world *ought* to be. That is, we want the world to conform more closely to our normative ideals, our mental models of how things should be. To say that something ought to be a particular way, implies that this imagined state manifests something *desirable, valuable*, and *worthy of endorsement*.[11] We could say that normative statements are intentional, rather than descriptive or explanatory. But unlike the methods of empirical analysis, that will be familiar to the reader,

7. Tufekci, "Opinion. We Already Know How to Protect Ourselves."
8. Letters.
9. Thomson.
10. Humberstone.
11. Gibbard; Korsgaard; Wedgwood.

we have far less understanding of what makes a good normative theory, because we are generally less familiar with normative analysis.

This essay will not focus on normative questions in general, but on a narrower set of questions concerning ideas of justice. Many normative questions concerning the actions of individuals or institutions are properly understood as ethical questions. For the purpose of this article I will define ethics as primarily concerned with individual decision making, with the question "How should I act?" and the attendant question, "What is involved in living a good life?" Justice involves not questions about individual actions (what should I do), or actions between individuals (how should I treat you, what do we owe each other), but with such questions as they arise across society. It is concerned with the principles for organizing wider social relations, what we owe one another across a society. Justice concerns our common life, how it should be ordered, what good government (and governance) looks like. Often questions of justice are concerned with the organization of the state. Yet, as our communication systems become ever more central to the structuring and organization of social and political relations, they too must become subject to considerations of justice.[12] In short, what I mean by justice is an ideal in which social, political, and indeed communicative arrangements are what they *ought* to be.[13] It concerns questions about what makes such arrangements good and right.

To illustrate the distinction between justice and ethics, consider the following propositions.

A: By right, you may insult Peter.
B: You should not insult Peter.

Both are normative statements in that they contain directives. They are statements about how the world *ought* to be. But we can interpret (A) as a statement about the requirements of justice and (B) as a statement about the requirements of ethical behavior. Not to insult people is a sound ethical principle (i.e., because it avoids rudeness and possible offence). It answers the question what one ought to do in a given situation, but is not necessarily a good principle for organizing wider social relations. In asking how we should act, we are not necessarily concerned with questions of justice. It is not a contradiction to say that while you shouldn't insult Peter, you nevertheless

12. See, for example, Fisher.
13. Michelman.

may. Affirming self-expression (i.e., because of its central importance to autonomy and the pursuit of truth), even of views we find distasteful, is a statement on how social and political arrangements across society ought to be organized. The distinction is important, because A and B clash. But principles of justice are of a higher order than those of ethical behavior—at least if we accept that human autonomy and truth should enjoy priority in our social ordering over the avoidance of rudeness and offence.[14] In this respect, an action can be consistent with justice, but unethical.

In exploring normative analysis, my primary aim is to expand the discussion about what makes good normative theories for the communications field. Normative work clearly has a strong appeal in the field. After all, media and communications scholarship is rich in critical evaluations. Evidently, many scholars are eager to address wrongs and injustices through their work. Sadly, critical appraisals outweigh positive visions of how we may move toward a more just state of affairs. Admittedly, it is easier to know what we are against, than to say precisely what we are for. That is why a secondary aim of this essay is to encourage the development of positive conceptions of what justice demands of communication and its technologies. By encouraging and broadening systematic engagement with normative analysis I hope to encourage and contribute to the development of systematic and rigorous conceptions of communication justice.

Section one of the essay outlines the basic elements that ideas of justice typically consist of. While the essay does not advance a particular conception of justice, it will endorse a republican conception of freedom as nondomination for illustrative purposes. Section two outlines a taxonomy that distinguishes between communication as a means to, and end of justice. It asks what gives certain communicative processes, practices, technologies, institutions, and other informational and communicative resources exceptional value, meriting special protection. I argue that this taxonomy is important if we are to approach questions of justice in a systematic and consistent manner within the communication field.

Goods, Principles, and Procedures

There are two justice-related questions that normative analysis tries to answer. Floyd refers to these as the organizing question: How should we

14. Ash.

arrange our sociopolitical order? and the foundational question: What justifies our answer?[15] Here, I will focus primarily on the organizing question. The question what our social, political (and indeed communicative) arrangements *ought* to be is, in the broadest sense, a question about what justice consists of. Most answers to this question involve stipulations about (a) essential or primary goods (things that are essential to everyone), (b) principles that tell us how these should be distributed, and/or (c) procedural principles that tell us how to decide matters of (a) and (b).

Primary Goods

Answering the organizing question always involves some "currency" in terms of which an idea of justice is articulated. Currencies are those essential, substantive, key, or primary goods (as Rawls[16] calls them in his *Theory of Justice*) that justice should secure. They are those things anyone would need to live a good life, the universal means needed for obtaining human ends. Different traditions of justice reach different assessments of the kind of things people have universal reasons to value. For Rawls,[17] these include economic resources (welfare, jobs, entitlements) and liberties. For Nozick[18] and Scanlon[19] rights and freedoms, a different conception of freedom for Pettit[20], and resources for Dworkin.[21] For Sen[22] and Nussbaum[23] the capabilities required to realize key functionings,[24] for Utilitarians like Singer, Mill, or Bentham welfare, pleasure, or happiness. All of these are goods that one can reasonably argue are needed for living a good life, be it some resources, freedoms, and rights, or the actual capability to realize a right or freedom. Put simply, saying that a certain set of arrangements is just, is to say that it distributes those goods that are needed to live well in the right way.

Primary goods also provide standards for appraising justice: are rights and freedoms obtained, functionings realized, welfare maximized, or

15. Floyd.
16. Rawls, *A Theory of Justice*.
17. Ibid.
18. Nozick.
19. Scanlon.
20. Pettit, *Republicanism.*
21. Dworkin.
22. Sen.
23. Nussbaum.
24. See also Schejter and Tirosh.

necessary resources available. Shifting the focus onto communicative goods, the question may be whether people enjoy the rights, resources, and/or capability to access quality information, or whether they are able to freely express legitimate grievances. While it is tempting to compile an extensive list of primary goods (who, after all, could disagree that all the goods listed above are important), considered parsimony makes for a better answer to the organizing question. After all, different goods can come into conflict with each other: ensuring everyone has sufficient resources may conflict with the right to private property, for instance. In the domain of communication, the right to free expression may conflict with the provision of quality information. To deal with such conflicts requires us to establish a priority between different goods, to decide which we have the most fundamental reasons to value.

Principles

The reader might have an inkling that an answer to the organizing question that merely stipulated a set of primary goods justice should secure would be incomplete. Justice, after all, is concerned with the obligations people have toward each other across society. We could stipulate that everyone should have the capability to express themselves, and the resources necessary to access quality information. However, if this is not already the case, stipulating that a certain communication right, resource or capability is a key good, is insufficient. We must also say how this good is to be allocated, and how the burden of its provision is to be distributed. Are individuals, families, or groups entitled to these goods? And is it local communities, states, technology, or media companies, the most advantaged, or everyone globally who is obliged to help provide them? Theories of justice differ not only on their list of primary goods, but also on the principles used to determine their allocation.

Part two of Rawl's[25] second principle of justice, known as the "difference principle" is one means for determining when a redistribution of some resource or good is permissible: namely, when it benefits the least advantaged members of society. In the context of freedom and equality of expression, the difference principle may guide us in considering questions of media ownership. For instance, that "an increased concentration in media ownership is permitted only if it will benefit the representation of

25. Rawls, *A Theory of Justice*.

the most marginalized voices in society, and hence can be seen to increase the pluralism of information and perspectives available in the media."[26] Contrast this with utilitarian theories of justice, for instance, whose central principle is to maximize the aggregate good (e.g., happiness, pleasure, voice). Utilitarian theories would thus permit a change in the concentration of media ownership or power if it increased aggregate representation of voice, even if it did not improve, or even worsened the situation of the least advantaged.

An answer to the organizing question that defines a fixed set of primary goods and invariable principles for their allocation is referred to as a substantive conception of justice, because its parameters fix the distribution of primary goods our social and political arrangements ought to secure. However, at least since Rawls, political theory has recognized social pluralism, the diversity of cultural values and ways of life (what Rawls calls "comprehensive doctrines"), as a key challenge for theories of justice. Rawls asked whether it was "possible for there to exist over time a just and stable society of free and equal citizens, who remain profoundly divided by reasonable religious, philosophical and moral doctrines?"[27] Pluralism is the hard problem bedeviling answers to the organizing question. Because providing a fixed definition of primary goods and their ideal distribution, also fixes what a good or just society should ultimately look like, it is a matter of controversy whether substantive conceptions of justice can truly accommodate social pluralism.

Pluralism exposes substantive conceptions to the critique of ethnocentrisms, for how can they justify universal value of their conception in culturally impartial ways? How can it be guaranteed that the substantive goods stipulated are similarly valued by different members of pluralistic societies? After all, primary goods (as universal means to human ends) and their distribution should be sufficiently general to allow all of us to value them independently of the specific lives we want to live, our cultural, or religious heritage. To address the challenge of pluralism most answers to the organizing question involve some procedural principles: Rather than providing a fixed definition of just social and political arrangements, they define procedures through which these arrangements should be decided. Procedural (rather than substantive) answers to the organizing question

26. Hänska; see also Chapter 3 in Schejter and Tirosh.

27. Rawls, *Political Liberalism*, 4.

want to remain sensitive to the pluralism of views in society by involving those who live within a sociopolitical order in affirming its parameters.

Procedures

Procedural conceptions of justice attempt solve the challenge of pluralism by avoiding *ex ante* substantive commitments to a particular combination of primary goods and principles. Instead they stipulate some kind of inclusive (imagined or real) process for deciding the contours of justice itself. Rather than offer a substantive answer to the question "What do just social arrangements require?", they provide us with a method for answering the question. In this sense, procedural conceptions are not directly concerned with primary goods and their allocation, but with legitimate and fair ways of deciding which primary goods should be distributed and how.[28] The advantage of procedural conceptions is that they are supposed to be more sensitive to pluralism. Incorporating the democratic axiom that those bound by the law should partake in shaping it, justice resides not in the ideal allocation of universally valued primary goods, but in the process through which we come to determine a specific allocation of goods. Procedures are open ended, such that the shape of just social arrangements may change through time and circumstance. What matters to justice in the procedural view is not, above all, whether a set of substantive conditions are met, but whether the process through which arrangements are decided and put in place meets a set of procedural conditions.

Most procedural theories build on some variant of the idea of public reason to specify the procedural standards that decisions should meet. Rawls, for instance, makes provisions for reasonableness of those engaged in the process of deciding the shape of justice, for the notional person behind his "veil of ignorance." Deliberative democrats circumscribe deliberation, and make requirements for inclusiveness, rationality, and the quality of argument, which aim to place constraints on the process of deciding just arrangements—those famous requirements often known as the "Habermasian public sphere." Agonistic pluralists counter that the requirements of deliberation are too onerous and restrictive, that they

28. In distinguishing procedural from substantive conceptions, sometimes a distinction is made between justice and legitimacy, where the former is a feature of sociopolitical arrangements, and the latter a feature of procedures. Here justice resides in the legitimacy of the process, and implies an obligation to accept its results. See Pettit, "Legitimacy and Justice in Republican Perspective."

sanitize debate by removing legitimate voices that do not meet its restrictive constraints.[29] While Nussbaum[30] stipulates a list of 11 central capabilities, Sen[31] declines to do so, favoring local deliberative processes for determining salient capabilities because deliberation is more sensitive to pluralism. Sen also deploys comparative approaches for identifying the most pressing injustices, rather than developing a comprehensive but abstract conception of perfect justice.

A Balance Between Procedures and Substantive Outcomes

Most conceptions of justice are a mixture of substance and procedure. Rawls[32] balances substantive stipulations with hypothetical procedures (the "veil of ignorance") for determining how predefined primary goods (resources and liberties) should be allocated. Deliberative democrats tilt the balance decisively toward procedures, with the aim of giving just social arrangement dialogic, intersubjective foundations, by emphasizing communicative processes for *deciding* what allocation of which goods justice requires.[33]

Procedures are (ostensibly) more sensitive to pluralism, because they seek everyone's input, they should be fair and not prejudice the outcome, thus making space for various comprehensive doctrines. Yet, we would also reject a procedure if its substantive outcome is perceived as unjust. Imagine an inclusive and fair deliberative process through which all men and women decided by consensus that all future decisions should be made only by women. This is why many theories of justice try to strike a balance between procedures and substantive conceptions, usually making provisions, which constrain procedures. As noted, for instance, deliberative democrats stipulate conditions of reasonableness and rationality, which agonistic democrats deem too restrictive. Liberal democracies employ institutions to constrain fickle and capricious shifts in public will.

The distinction between substantive and procedural conceptions of justice bears emphasizing, because communication can play a critical role in both. Liberty is hard to conceive without the right to free expression—as such communication can be critical to substantive conceptions of justice.

---

29. See Mouffe.
30. Nussbaum.
31. Sen.
32. Rawls, *A Theory of Justice*.
33. Chambers.

But expression, quality of information, debate, and deliberation are also the lifeblood of procedural conceptions of justice.

Organizing Ideal

Earlier on I suggested, following Floyd, that normative analysis seeks to answer two key questions: the organizing question (how should we arrange our sociopolitical order) and the foundational question (what justifies our answer). Whether we find the justification for a particular conception of justice compelling—whether normative analysis sways us—often depends on whether we are already committed to some underlying value. I call these underlying values organizing ideals, but we could also refer to them as primary political values. They are, in a sense, the bedrock of justice.

Organizing ideals undergird every conception of justice, sometimes lurking in the background without being explicitly articulated. We cannot make sense of deliberative ideals, unless we recognize that they rest on some prior commitment to reason and autonomy. Rawls's conception of justice rests on some prior commitment to freedom and equality. Sen's and Nussbaum's on a commitment to human flourishing. There are many possible organizing ideals, including individual freedom (understood as nonintervention or nondomination), autonomy, reason, and equality. Similarly, general welfare, community, flourishing, or happiness are all potential organizing ideals. To endorse one of these is not to say that others are unimportant, but to establish a priority between them, to say which is a primary, and which a subordinate value. For example, we may value equality, but its appeal may be based on some prior commitment to nondomination, in that the value of equality derives from a prior commitment to freedom from domination. Clarity about the organizing ideal we are committed to helps us work out which primary goods, principles, and procedures are consistent with and conducive to this ideal. We could say that primary goods, principles, and procedures are constitutive of the organizing ideal, things that are substantively part of what it means to be free or equal, for example. It is worth noting that procedural conceptions encounter a problem of circularity here: On the one hand, procedures are supposed to be more sensitive to pluralism because they avoid prior commitments to fixed ideals, yet the parameters we stipulate for procedures (e.g., reasonableness, equal participation) can only be justified by appeal to some organizing ideal. In Box 1, I sketch a communicative conception of justice based on the organizing ideal of freedom as nondomination.

Box 1 Freedom as Nondomination

The aim of this essay is to explore normative analysis in the communication field. To put flesh on the subject matter, let me commit to an organizing ideal for illustrative purposes. Though I find this ideal personally compelling, it is not necessary for the reader to share this commitment. Let us then take freedom as the most basic political value, as our organizing ideal. More specifically, let us commit to a notion of freedom understood as nondomination rather than noninterference.[34] For it is conceivable to be free of external interference but still subject to domination, in that the freedom enjoyed depends on the largesse of others (e.g., a benign dictator) who may rescind this freedom at any time. In other words, we may be free but at risk of arbitrary interference. Consider the freedom to communicate that social media platforms afford. Yet these platforms also offer expansive opportunities for control and domination to those who control the social graph. Our communications are free (in a sense) but we are at risk of uncontrolled and arbitrary interference from others, and thus subject to domination. Note also, that a commitment to freedom as nondomination as organizing ideal does not discount other values such as equality or flourishing, but that it establishes a priority. If and when we face trade-offs between freedom and other values, freedom has priority.

Our tentative answer to the organizing question could place the onus on proceduralism, in that social and political arrangements free of domination are best achieved through processes of collective choice that remain open ended (where every choice is temporary, and subject to potential revision in the future). Choice includes important communicative elements. Let us suggest two key communicative goods that justice should secure: (1) the capability to express legitimate grievances and have these recognized and (2) the resource of quality information. The continuous presence of countervailing voices is critical to prevent domination, and quality information is essential to human autonomy. Theses goods are to be allocated in such a manner that they have the greatest benefit for those most vulnerable to domination. This sketch of a normative analysis stipulates two communicative goods, a principle for their allocation, conceived as procedural constituents of the organizing ideal of freedom as nondomination.

34. Pettit, *Republicanism*; Shapiro.

Communicative Means and Ends of Justice

So far I have argued that we should attune ourselves to the nature of normative analysis. That when we make normative claims these are either evaluative or directive, statements about how things ought to be. And that justice concerns the organizing question: how sociopolitical relations should be organized more generally, not questions that pertain only to individual actions. Answering the organizing question usually involves a combination of stipulating primary goods that justice should secure, principles governing their allocation, and (or sometimes alternatively) procedures to help decide what kind of goods justice should secure and how it should distribute these.

The aim here is to locate communication in ideas of justice. The key question is how communication is related to sources of value: namely, to organizing ideals, primary goods, principals, and procedures. How do certain communicative processes, practices (capabilities), technologies, institutions, and other informational and communicative resources derive value? I propose that they can derive value in one of three ways: Communication can be constitutive of justice (an end of justice), and intrinsically valuable, either as (1) substantive good, or because of its (2) procedural value. Communication can also be (3) a means to justice, in that it can be causally necessary for us to obtain some primary or procedural good, and thus extrinsically valuable (see Table 1).

Communication as a Substantive End of Justice

Specific communicative processes, practices (capabilities), technologies, institutions, and other informational and communicative resources can be intrinsically valuable. When communication has intrinsic value, this

TABLE 1 How communication can relate to justice

	Procedural Justice	**Substantive Justice**
Ends/Constitutive value	Communication as key to procedural conceptions of justice, for instance, in deliberation. Here procedures and processes are seen as constitutive of the organizing ideal, and the locus of justice itself.	Some communication practices, capabilities (e.g., expression), or resources (e.g., quality information) can be viewed as substantive goods in their own right.
Means/Derivative value	Communication as a means of achieving justice, as central to nonideal theories of justice.	

is either because it is a substantive good in its own right, or because it is central to a procedural conception of justice. In either case, communication is constitutive of an organizing ideal, and can therefore be understood as an end of justice itself.

Substantive conceptions of justice, as outlined above, define primary goods that we should all have reasons to value, as the grounds on which justice is assessed. Depending on our answer to the organizing question, some communication rights or capabilities could be conceived as such primary goods. We can value self-expression and voice as primary goods,[35] for instance, because expression can be viewed as the very substance of freedom. Similarly, we could include access to quality information, or the right to privacy, as the most basic goods that anyone would need to live a flourishing life. Though expression, voice, privacy, and information by no means exhaust the list of possible primary goods (depending on the underlying organizing ideal to which we are committed) they could all be said to have intrinsic normative value. That is, we can argue that they are not merely related to, derivative from, or supportive of justice, but substantially constitutive of some organizing ideal (e.g., freedom, flourishing, general welfare). Just sociopolitical arrangements thus prevail to the extent that a fair distribution of these goods is secured.

Communication as Procedural End of Justice

A procedure, rather than as a specific state of affairs, could also be our answer to the organizing question. Here, key communicative practices (capabilities), technologies, institutions, and other informational and communicative resources have substantive procedural value (because they are important to the legitimacy of deciding on sociopolitical arrangements). For deliberative democrats, for example, communication is critical for reaching a consensus, accommodating differences, and facilitating a process of transforming people's preferences through rational debate. For Sen, deliberation is the process through which communities determine locally critical capabilities, their provision, and distribution. Many procedural conceptions of justice build on the organizing ideal of public reason, which is supposed to be engendered in public discourse and debate. Communication is intrinsic to the exercise of public reason. And rational and inclusive collective choices (through which the parameters of social order are to be determined)

35. Ash; Couldry; see Table 2.

are considered to be the very locus of justice. Rather than providing a substantive answer to the organizing questions, in proceduralism the justness of social-political arrangements depends on them being open to revision through fair and legitimate processes of collective choice.

Consider the organizing ideal of freedom as nondomination (see Boxes 1–3). Because most societies are pluralistic, a substantial risk of injustice emerges from the possibility of some group losing its autonomy as another group becomes dominant and imposes its preferred order on everyone else—in which case one group would get what it wants, but all others would be subject to uncontrolled interference. Pluralism of competing forces countervails the risk of any one group becoming dominant. Expressing ones views and interests, and having these recognized, is a desideratum of such a countervailing force, and can be viewed as constitutive of political arrangements free of domination. Repurposing Rawls's difference principle we may suggest that the voices of those most vulnerable to domination should enjoy priority, for instance. Here, the capability to express legitimate grievances is valued intrinsically because it constitutes the substance of nondomination (see Box 2). It is a procedural good central to the legitimacy of the processes in which justice resides.[36] Deliberativeness, reasonableness, or participation could also be procedural goods, the absence of which would signal a certain injustice or illegitimacy of prevailing social and political arrangements.

Box 2 A Procedural Conception of Nondomination

In the procedural conception of freedom as nondomination outlined in Box 1, domination is countervailed by fostering two communicative goods: (1) The capability to express legitimate grievances and have these recognized and (2) the provision of quality information. In this view, quality information is a substantive good, because the autonomy that nondomination should secure could not be realized without reliable information on which to base individual choices. The practices and institutions that provide quality information have derivative (extrinsic) value. The capability to express legitimate grievances and have these recognized is a procedural good (with intrinsic value), in that it is constitutive of arrangements that are free of domination. It countervails domination, and serves as an indicator of freedom.

36. Besley and McComas; Schaefer.

Communication as a Means to Justice

Normative analysis should not only be concerned with answering the organizing question, but also with how an idea of justice can be realized. Rawls famously distinguished between ideal and nonideal theory.[37] Ideal theory concerns the definition of what justice consists in, and the justification of that definition. Nonideal theory concerns the realization of justice. This three-way distinction—between nonideal and ideal theories, and within ideal theory between substantive and procedural conceptions—augments our understanding of what a just society demands of communication. While communicative processes, practices, technologies, institutions, and other informational and communicative resources can be ends of justice (i.e., constitutive of substantive or procedural conceptions of justice), they can also be a means to justice. They can be important components of nonideal theories.

There is a critical difference: if the capability to express legitimate grievances is constitutive of justice, it is sacrosanct. If, however, it is a means to justice (see Table 2), then it is not sacrosanct. Let us assume our organizing ideal is equality of welfare, then affording someone a voice can be a means to equality, in that it can help us detect salient inequalities. But the capability of expression would only be valuable insofar as it helps ensure a more equal distribution of welfare. Here, capability of expression itself is not constitutive of our idea of justice, and does not necessarily guarantee it. Things change, of course, if it is equality of voice that we are aiming for. Yet, if we are seeking equality of voice, then the

TABLE 2 The capability of expression as a means to, or end of justice

Free expression as substantive end	The capability of expression has intrinsic value, as a substantive part of justice.
Free expression as a procedural end	The capability of expression is an important part of fair and legitimate processes of collective choice, which themselves constitute nondomination. As such, voice has intrinsic procedural value.
Free expression as means	The capability of expression as means to secure equal distribution of welfare, for example. Some also argue that inclusive deliberation will help reach consensus, while others argue that it may actually increase disengagement. Insofar as it secures some of these goods it has extrinsic (secondary or derivative) value.

37. Simmons.

capability of expression may not suffice either. Consider the following: social media may afford everyone a nominal voice, yet that voice may well go unheard. On social media it is more effective to drown out voices you find undesirable (e.g., using botnets and sock-puppet accounts to flood platforms with preferred voices), rather than to censor and suppress them. Perhaps the thing we are seeking to secure is not expression as such, but having legitimate grievances acknowledged—in which case the capability of expression may be a necessary (but not necessarily sufficient) means for obtaining our ends.

Consider another example, the provision of high-quality information, and the role of journalism, media institutions, and technology platforms in securing access thereto. Is it high-quality information as such that is valued, or does it serve as a proxy for political knowledge? Political knowledge is generally regarded essential to the proper functioning of democracy.[38] Floridi's[39] information ethics is premised on the value of quality information to ethical decision making. But journalism may not necessarily ensure political knowledge. The recent abundance of what Wardle and Derakhshan[40] call "information disorder"—the proliferation of mis- and disinformation—reveals a tension between the means (provision of quality information) and the ends (political knowledge). We may not suffer a paucity of high-quality information, but an overwhelming volume of low-quality, even disinformation, that muddies the water, drowning-out high-quality information. Assuming this analysis is accurate, lack of political knowledge may not be caused by the absence of quality information, but by the prominence and cacophony of bad information. Therefore, increasing the provision of high-quality information may not necessarily be the most effective *means* of ensuring a well-informed citizenry.

A final example includes some strands of deliberative democracy, which hold that the aim of deliberative communication is to narrow disagreements, facilitate learning, transform preferences, or even mitigate cognitive biases.[41] Here deliberative communication has extrinsic value, derived from its ability to change minds, improve political knowledge, and facilitate decision making by bridging differences, converging opinions, and facilitating agreement. Often inclusion of a diversity of voices

38. Nielsen; Schudson.

39. Floridi.

40. Wardle and Derakhshan.

41. Dryzek, *Deliberative Democracy and Beyond*; "Deliberative Democracy in Divided Societies."

is considered critical to deliberation, yet as Mutz's[42] seminal work has shown, people often tend to disengage from processes the more they are exposed to different opinions. Ensuring the greatest diversity of voices may not necessarily facilitate agreement, but rather increase disengagement. Similarly, we may stipulate maximum inclusion to increase the subjective legitimacy of collective decision-making processes (i.e., whether a decision is perceived as legitimate by participants). But an inclusive, participatory choice process will not necessarily yield greater perceived legitimacy than a less participatory process. Britain's EU referendum was more participatory than a parliamentary decision, yet a decision by parliamentary delegates could plausibly have enjoyed greater (cross-cutting) perceived legitimacy than the referendum outcome does.

The distinction between means and ends may seem immaterial, but it has profound consequences for our thinking about the place of communicative processes, practices, technologies, institutions, and other informational and communicative resources in conceptions of justice, and for understanding what justice requires of communication. When communication is regarded as a substantive end or constitutive of some procedural conception of justice, the issue is quite simple. The absence of relevant communicative processes, capabilities, or other informational and communicative resources would indicate a degree of injustice. However, things get more complicated when we think of communication as a means to justice, where its presence or absence serves at best as a proxy measure for evaluating justice.

When the value of communications derives from some other good it can help us obtain (such as autonomy, political knowledge, equality, agreement, or subjective legitimacy), it is not intrinsically valuable, but valued as a means. But means are not sacrosanct. They are potentially fungible, substitutable without detrimentally affecting the overall justness of social arrangements. Constitutive elements of justice (its ends) are sacrosanct, and cannot be substituted. For instance, journalism may be sufficient but unnecessary for securing quality information (making it a means). And quality information may be necessary but insufficient for ensuring political knowledge (making it an end). Insofar as something other than good journalism could help obtain quality information, justice would not suffer the absence of journalism. As political knowledge is impossible without quality information, its absence would detrimentally

42. Mutz.

Box 3 The Means and Ends to Freedom as Nondomination

Journalism has extrinsic value as a sufficient but unnecessary means to quality information. As such, journalism and its institutions are potentially substitutable. The capability to express legitimate grievances and have these recognized is necessary but not necessarily sufficient for insuring nondomination. As such expression is an end of justice with substantive procedural value.

affect justice. In this view, journalism's value is derivative, but that of quality information is constitutive. Similarly, agreement between members of society may be a desideratum of procedural conceptions of justice, but maximizing the inclusion of voices may not actually be the best way of facilitating agreement. Insofar as something other than the widest inclusion of voices could help obtain an agreement, justice would not suffer the absence of some voices. However, hearing and acknowledging the diversity of voices present in society (particularly of its most marginalized members), is necessary for ensuring nondomination (see Box 3).

Normative analysis also requires us to carefully parse the trade-offs involved in prioritizing different organizing ideals and their constitutive goods, all of which we have significant reasons to value. This is an important task, especially if we want to offer a positive vision of communication justice that articulates what we are for (rather than limiting ourselves to critical assessments of those things we are against). For instance, we can value both the capability of expression and quality information, but under certain circumstances the two can compete and even collide—ensuring high-quality information will require us to privilege some voices (e.g., those of experts) over others. As we have seen, inclusive participation and subjective legitimacy can also become rivalrous: After all, more inclusive participation may reduce perceived legitimacy and engagement. On the other hand, a political decision may be widely viewed as legitimate, even if it was not particularly inclusive and participatory.

If our thinking on the place of communication in ideas of justice is to hit the ground of practice, clarity is needed on the priorities that different communicative practices and resources enjoy. How to prioritize voices if not all can be heard? How to prioritize between expression and quality information? Prioritizing requires clarity about the organizing ideal we endorse, from which the value of primary goods, principles, and procedures

ultimately derives. It requires clarity on whether the communicative practice or resource in question has intrinsic or extrinsic value. Properly parsing such trade-offs, under full considerations of competing values and the reasons lending them support, is what normative analysis is about.

Concluding Remarks

In this essay, I have been concerned with normative questions, more specifically questions about justice in the communications field. I have suggested that justice concerns the question how social and political relations should be organized—and the role communications ought to play in this organization. I have not attempted to offer a direct answer, but rather to shed some light on what is involved in answering such questions. If we are to think about what justice demands of communication, and to formulate robust normative views about media and communications, then we are engaged in the underappreciated enterprise of normative analysis. The fact that much social research takes its cue from a certain value-orientation, only underscores its relevance. Moreover, if we conclude empirical work with reflections on the broader significance of our findings, with an assessment of what is good, bad, or valuable about the phenomena we study, we should better be clear about how communication derives value.[43] We should know not only what we are against, but also articulate clearly what we are for. To that end, I hope this article can provide an impetus. That it can help us to understand the trade-offs and contradictions between different ways of valuing communication, that become essential in making informed assessments about what is good and bad about various communicative processes, practices, technologies, institutions, and other informational and communicative resources.

I want to end with a call for parsimony, that we carefully consider which specific communicative phenomena can (should) be usefully parsed in the language of justice, and why. Certainly, many communicative phenomena raise salient ethical considerations (questions about how they impact the quality of people's lives), but do not necessarily have any immediate purchase on questions of justice, because they are not immediately salient to wider questions about how a society ought to arrange its social and political affairs. The privacy of family photos shared online, or the kind of content

43. Althaus.

suitable to be aired pre- and post-watershed certainly raise ethical issues. But they do not have any immediate purchase on questions of justice. Not all communication is "justice apt." That is, not all communication is pertinent to questions of justice, and we must avoid conceptual overstretch, the familiar effort to shoehorn phenomena into a particular conceptual space where all kinds of things are to be parsed in terms of justice. At the same time, just because something is not justice apt, does not mean it is not valuable. But insofar as we believe that communication is justice salient, we need to know why. And knowing why requires us to understand what the underlying values are that determine communication's relationship to justice. Does communication have intrinsic value as a constituent of justice? Or, does it have extrinsic value because of its relationship to some primary good that, in turn, is a constituent of justice?

Finally, no communication right, resource or capability alone will be sufficient for justice to obtain—it would be absurd to claim that arranging our communicative processes, practices, technologies, institutions, and other informational and communicative resources in the right way is all that justice required. Nonetheless, communication can be an important component of justice, and under conditions of pluralism it is hard to imagine how justice could be realized without communications that help mediate between society's various interests, views, and ethical outlooks. It is also hard to imagine how people can become autonomous absent high-quality information to base their choices on—they may still be making their own choices, but those choices would less likely deliver the intended outcome. Good communication does not guarantee justice, but under conditions of pluralism it is essential and necessary for its pursuit.

BIBLIOGRAPHY

Althaus, Scott L. "What's Good and Bad in Political Communication Research? Normative Standards for Evaluating Media and Citizen Performance." In *Sage Handbook of Political Communication*, edited by Holli Semetko and Margaret Scammell, 97–111. London: Sage, 2011.

Ash, Timothy Garton. *Free Speech: Ten Principles for a Connected World*. London: Atlantic Books, 2016.

Besley, John C., and Katherine A. McComas. "Framing Justice: Using the Concept of Procedural Justice to Advance Political Communication Research." *Communication Theory* 15, no. 4 (November 1, 2005): 414–36. doi:10.1111/j.1468-2885.2005.tb00342.x.

Bucy, Erik P., and Paul D'Angelo. "Democratic Realism, Neoconservatism, and the Normative Underpinnings of Political Communication Research." *Mass Communication and Society* 7, no. 1 (March 1, 2004): 3–28. doi:10.1207/s15327825mcs0701_2.

Chambers, Simone. "Deliberative Democratic Theory." *Annual Review of Political Science* 6 (2003): 307–26.

Couldry, Nick. *Why Voice Matters: Culture and Politics after Neoliberalism*. London: Sage, 2010.

Dryzek, John. *Deliberative Democracy and Beyond: Liberals, Critics, Contestation*. Oxford: Oxford University Press, 2002.

———. "Deliberative Democracy in Divided Societies: Alternatives to Agonism and Analgesia." *Political Theory* 33, no. 2 (2005): 218–42. doi:10.1177/0090591704268372.

Dworkin, Ronald. *Taking Rights Seriously*. Vol. 136. Cambridge, MA: Harvard University Press, 1978.

Fisher, Max. "Inside Facebook's Secret Rulebook for Global Political Speech." *The New York Times*, December 27, 2018, sec. World. Accessed December 28, 2018. https://www.nytimes.com/2018/12/27/world/facebook-moderators.html.

Floridi, Luciano. *The Ethics of Information*. Oxford: Oxford University Press, 2013.

Floyd, Jonathan. *Is Political Philosophy Impossible?: Thoughts and Behaviour in Normative Political Theory*. Cambridge, UK: Cambridge University Press, 2017.

Gerhards, Jürgen, and Friedhelm Neidhardt. "Strukturen Und Funktionen Moderner Öffentlichkeit. Fragestellungen Und Ansätze." In *Politische Kommunikation: Grundlagen, Strukturen, Prozesse*, edited by Wolfgang R. Langenbucher, 52–88. Vienna: Braumüller, 1993.

Gibbard, Allan. *Meaning and Normativity*. Oxford: Oxford University Press, 2012.

Habermas, Jürgen. *The Structural Transformation of the Public Sphere: An Inquiry Into a Category of Bourgeois Society*. Cambridge, UK: Polity Press, 1992.

Hall, Stuart, and Henry Louis Gates. *The Fateful Triangle: Race, Ethnicity, Nation*. Cambridge, MA: Harvard University Press, 2017.

Hänska, Max. "Media Pluralism: How Rawls Can Help Us Think about Newscorp's BskyB Bid." *Polis: Journalism and Society* (blog), January 18, 2011. Accessed May 15, 2016. http://blogs.lse.ac.uk/polis/2011/01/18/media-pluralism-how-rawls-can-help-us-think-about-newscorp's-bskyb-bid-guest-blog/.

Harvey, Alison, and Koen Leurs. "Networked (in) Justice: An Introduction to the #AoIR17 Special Issue." *Information, Communication & Society* 21, no. 6 (June 3, 2018): 793–801. doi:10.1080/1369118X.2018.1438493.

Humberstone, I. L. "Direction of Fit." *Mind* 101, no. 401 (January 1, 1992): 59–84. doi:10.1093/mind/101.401.59.

Korsgaard, Christine M. *The Sources of Normativity*. Cambridge, UK: Cambridge University Press, 1996.

Letters. "Stop Jeremy Corbyn's Mistrial by Media over Antisemitism | Letters." *The Guardian*, April 2, 2018, sec. Politics. Accessed April 28, 2018. http://www.theguardian.com/politics/2018/apr/02/stop-jeremy-corbyns-trial-by-media-over-antisemitism.

Michelman, Frank I. "How Can the People Ever Make the Laws? A Critique of Deliberative Democracy." In *Deliberative Democracy: Essays on Reason and Politics*, edited by James Bohman and William Rehg, 145–72. Cambridge, MA: MIT Press, 1997.

Mouffe, Chantal. "Deliberative Democracy or Agonistic Pluralism." *Social Research* 66, no. 3 (1999): 745–58.

Mutz, Diana C. *Hearing the Other Side: Deliberative Versus Participatory Democracy*. Cambridge, UK: Cambridge University Press, 2006.

Nielsen, Rasmus Kleis. "The One Thing Journalism Just Might Do for Democracy." *Journalism Studies* 18, no. 10 (June 22, 2017): 1251–62. doi:10.1080/1461670X.2017.1338152.

Noelle-Neumann, Elisabeth. "The Spiral of Silence a Theory of Public Opinion." *Journal of Communication* 24, no. 2 (1974): 43–51. doi:10.1111/j.1460-2466.1974.tb00367.x.

Nozick, Robert. *Anarchy, State, and Utopia*. Vol. 5038. New York: Basic Books, 1974.

Nussbaum, Martha C. *Creating Capabilities*. Cambridge, MA: Harvard University Press, 2011.

Pettit, Philip. *Republicanism: A Theory of Freedom and Government*. Oxford Political Theory. Oxford, New York: Clarendon Press; Oxford University Press, 1997.

———. "Legitimacy and Justice in Republican Perspective." *Current Legal Problems* 65 (2012), 59–82.

Rawls, John. *A Theory of Justice*. Cambridge: Belknap Press, 1971.

———. *Political Liberalism*. New York: Columbia University Press, 1993.

Scanlon, T. M. *The Difficulty of Tolerance: Essays in Political Philosophy*. Cambridge, UK: Cambridge University Press, 2003.

Schaefer, David Lewis. "Procedural versus Substantive Justice: Rawls and Nozick." *Social Philosophy and Policy* 24, no. 1 (2007): 164–86. doi:10.1017/S0265052507070070.

Schejter, Amit M., and Noam Tirosh. *A Justice-Based Approach for New Media Policy: In the Paths of Righteousness*. New York: Springer, 2016.

Schudson, Michael. *Why Democracies Need an Unlovable Press*. Malden: John Wiley, 2013.

Sen, Amartya. *The Idea of Justice*. Cambridge, MA: Belknap, 2009.

Shapiro, Ian. "On Non-Domination." *University of Toronto Law Journal 62*, no.3 (July 5, 2012): 293-335. doi.10.3138/utlj.62.3.293.

Simmons, A. John. "Ideal and Nonideal Theory." *Philosophy & Public Affairs* 38, no. 1 (January 1, 2010): 5–36. doi:10.1111/j.1088-4963.2009.01172.x.

Taylor, Linnet. "What Is Data Justice? The Case for Connecting Digital Rights and Freedoms Globally." *Big Data & Society* 4, no. 2 (December 1, 2017): 1–14. doi:10.1177/2053951717736335.

Thomson, Judith Jarvis. *Normativity*. Peru, IL: Open Court, 2008.

Tufekci, Zeynep. "Opinion. We Already Know How to Protect Ourselves From Facebook." *The New York Times*, April 11, 2018, sec. Opinion. Accessed April 28, 2018. https://www.nytimes.com/2018/04/09/opinion/zuckerberg-testify-congress.html.

———. *Twitter and Tear Gas: The Power and Fragility of Networked Protest*. New Haven: Yale University Press, 2017.

Wardle, Claire, and Hossein Derakhshan. "Information Disorder: Toward an Interdisciplinary Framework for Research and Policymaking." Council of Europe, 2017. (Accessed November 1, 2017. https://shorensteincenter.org/information-disorder-framework-for-research-and-policymaking/.

Wedgwood, Ralph. *The Nature of Normativity*. Oxford: Oxford University Press, 2007.

Normativity

Normative generally means relating to an evaluative standard. **Normativity** is the phenomenon in human societies of designating some actions or outcomes as good or desirable or permissible and others as bad or undesirable or impermissible. A norm in this normative sense means a standard for evaluating or making judgments about behavior or outcomes. Normative is sometimes also used, somewhat confusingly, to mean relating to a descriptive standard: doing what is normally done or what most others are expected to do in practice. In this sense a norm is not evaluative, a basis for judging behavior or outcomes; it is simply a fact or observation about behavior or outcomes, without judgment. Many researchers in science, law, and philosophy try to restrict the use of the term normative to the evaluative sense and refer to the description of behavior and outcomes as positive, descriptive, predictive, or empirical.[1][2]

Normative has specialised meanings in different academic disciplines such as philosophy, social sciences, and law. In most contexts, normative means 'relating to an evaluation or value judgment.' Normative propositions tend to evaluate some object or some course of action. Normative content differs from descriptive content.[3]

One of the major developments in analytic philosophy has seen the reach of normativity spread to virtually all corners of the field, from ethics and the philosophy of action, to epistemology, metaphysics, and the philosophy of science. Saul Kripke famously showed that rules (including mathematical rules, such as the repetition of a decimal pattern) are normative in an important respect.[4][5]

Though philosophers disagree about how normativity should be understood, it has become increasingly common to understand normative claims as claims about reasons.[4] As Derek Parfit explains:

> We can have reasons to believe something, to do something, to have some desire or aim, and to have many other attitudes and emotions, such as fear, regret, and hope. Reasons are given by facts, such as the fact that someone's finger-prints are on some gun, or that calling an ambulance would save someone's life. It is hard to explain the *concept* of a reason, or what the phrase 'a reason' means. Facts give us reasons, we might say, when they count in favour of our having some attitude, or our acting in some way. But 'counts in favour of' means roughly 'gives a reason for'. The concept of a reason is best explained by example. One example is the thought that we always have a reason to want to avoid being in agony.[6]

Contents

Further reading

Philosophy

In philosophy, normative theory aims to make moral judgements on events, focusing on preserving something they deem as morally good, or preventing a change for the worse.[7] The theory has its origins in Greece.[8] Normative statements make claims about how institutions should or ought to be designed, how to value them, which things are good or bad, and which actions are right or wrong.[9] Normative claims are usually contrasted with positive (i.e. descriptive, explanatory, or constative) claims when describing types of theories, beliefs, or propositions. Positive statements are (purportedly) factual, empirical statements that attempt to describe reality.

For example, "children should eat vegetables", and "those who would sacrifice liberty for security deserve neither" are normative claims. On the other hand, "vegetables contain a relatively high proportion of vitamins", and "a common consequence of sacrificing liberty for security is a loss of both" are positive claims. Whether a statement is normative is logically independent of whether it is verified, verifiable, or popularly held.

There are several schools of thought regarding the status of normative statements and whether they can be rationally discussed or defended. Among these schools are the tradition of practical reason extending from Aristotle through Kant to Habermas, which asserts that they can, and the tradition of emotivism, which maintains that they are merely expressions of emotions and have no cognitive content.

There is large debate in philosophy surrounding the normative and whether you can get a normative statement from an empirical one (ie whether you can get an 'ought' from an 'is', or a 'value' from a 'fact'). Aristotle is one scholar who believed that you could in fact get an ought from an is. He believed that the universe was teleological and that everything in it has a purpose. To explain why something is a certain way, Aristotle believed you could simply say that it is trying to be what it ought to be.[10] On the contrary, David Hume believed you cannot get an ought from an is because no matter how much you think something ought to be a certain way it will not change the way it is. Despite this, Hume used empirical experimental methods whilst looking at the normative. Similar to this was Kames, who also used the study of facts and objective to discover a correct system of morals.[11] The assumption that 'is' can lead to 'ought' is an important component of the philosophy of Roy Bhaskar.[12]

Normative statements and norms, as well as their meanings, are an integral part of human life. They are fundamental for prioritizing goals and organizing and planning. Thought, belief, emotion, and action are the basis of much ethical and political discourse; indeed, normativity is arguably the key feature distinguishing ethical and political discourse from other discourses (such as natural science).

Much modern moral/ethical philosophy takes as its starting point the apparent variance between peoples and cultures regarding the ways they define what is considered to be appropriate/desirable/praiseworthy/valuable/good etc. (In other words, variance in how individuals, groups and societies define what is in accordance with their normative standards.) This has led philosophers such as A.J. Ayer and J.L. Mackie (for different reasons and in different ways) to cast doubt on the meaningfulness of normative statements. However, other philosophers, such as Christine Korsgaard, have argued for a source of normative value which is independent of individuals' subjective morality and which consequently attains (a lesser or greater degree of) objectivity.[13]

Social sciences

In the social sciences, the term "normative" has broadly the same meaning as its usage in philosophy, but may also relate, in a sociological context, to the role of cultural 'norms'; the shared values or institutions that structural functionalists regard as constitutive of the social structure and social cohesion. These values and units of socialization thus act to encourage or enforce social activity and outcomes that *ought* to (with respect to the norms implicit in those structures) occur, while discouraging or preventing social activity that *ought not* occur. That is, they promote social activity that is socially *valued* (see philosophy above). While there are always anomalies in social activity (typically described as "crime" or anti-social behaviour, see also normality (behavior)) the normative effects of popularly endorsed beliefs (such as "family values" or "common sense") push most social activity towards a generally homogeneous set. From such reasoning, however, functionalism shares an affinity with ideological conservatism.

Normative economics deals with questions of what sort of economic policies should be pursued, in order to achieve desired (that is, valued) economic outcomes.

International relations

In the academic discipline of International relations, Smith, Baylis & Owens in the *Introduction* to their 2008 [14] book make the case that the normative position or normative theory is to make the world a better place and that this theoretical worldview aims to do so by being aware of implicit assumptions and explicit assumptions that constitute a non-normative position, and align or position the normative towards the loci of other key socio-political theories such as political liberalism, Marxism, political constructivism, political realism, political idealism and political globalization.

Law

In law, as an academic discipline, the term "normative" is used to describe the way something ought to be done according to a value position. As such, normative arguments can be conflicting, insofar as different values can be inconsistent with one another. For example, from one normative value position the purpose of the criminal process may be to repress crime. From another value position, the purpose of the criminal justice system could be to protect individuals from the moral harm of wrongful conviction.

Standards documents

Normative elements are defined in International Organization for Standardization Directives Part 2 as "elements that describe the scope of the document, and which set out provisions". Provisions include "requirements", "recommendations" and "statements". "Statements" include permissions, possibilities and capabilities. A "requirement" is an "expression in the content of a document conveying criteria to be fulfilled if compliance with the document is to be claimed and from which no deviation is permitted." It is not necessary to comply with recommendations and statements in order to comply with the standard; it is necessary to comply only with the requirements (that are denoted by the verbal form "shall"). There is much confusion between "normative" and "requirement", however the ISO terminology is supported by national standards bodies worldwide and is the legitimate description of these terms in the context of standards documents.

In standards terminology still used by some organisations, "normative" means "considered to be a prescriptive part of the standard". It characterises that part of the standard which describes what *ought* (see philosophy above) to be done within the application of that standard. It is implicit that application of that standard will result in a *valuable* outcome (*ibid.*). For example, many standards have an introduction, preface, or summary that is considered non-normative, as well as a main body that is considered normative.

"Compliance" is defined as "complies with the normative sections of the standard"; an object that complies with the normative sections but not the non-normative sections of a standard is still considered to be in compliance.

- Normative = prescriptive = how to comply
- Informative = descriptive = help with conceptual understanding

Typically, *normative* is contrasted with *informative* (referring to the standard's descriptive, explanatory or positive content). Informative data is supplemental information such as additional guidance, supplemental recommendations, tutorials, commentary as well as background, history, development, and relationship with other elements. Informative data is not a requirement and doesn't compel compliance.

See also

- Conformity
- Decision theory
- Economics
- Hypothesis
- Is-ought problem
- Linguistic prescription
- Norm (philosophy)
- Normative economics
- Normative ethics
- Normative science
- Philosophy of law
- Political science
- Scientific method
- Value

References

1. Bicchieri, Cristina (2005). *The Grammar of Society:The Nature and Dynamics of Social Norms*. Cambridge University Press. ISBN 978-0521574907.
2. Bicchieri, Cristina (2017). *Norms in the Wild: How to Diagnose, Measure, and Change Social Norms*. Oxford University Press. ISBN 9780190622053.
3. Jarvis., Thomson, Judith (2008). *Normativity*. Chicago, Ill.: Open Court. ISBN 9780812696585. OCLC 227918828 (https://www.worldcat.org/oclc/227918828).
4. Thomas, Scanlon (2014). *Being realistic about reasons* (First ed.). Oxford. ISBN 9780199678488. OCLC 862091562 (https://www.worldcat.org/oclc/862091562).
5. 1940-, Kripke, Saul A. (1982). *Wittgenstein on rules and private language : an elementary exposition*. Cambridge, Mass.: Harvard University Press. ISBN 0674954009. OCLC 7998796 (https://www.worldcat.org/oclc/7998796).
6. Parfit, Derek (2011). *On what matters Volume one* (https://archive.org/details/onwhatmattersvol0001parf). Edited and Introduced by Samuel Scheffler. Oxford: Oxford University Press. ISBN 9780191576706. OCLC 744616054 (https://www.worldcat.org/oclc/744616054).
7. Strauss, Leo (2017-09-08), "What Is Political Philosophy? The Problem of Political Philosophy" (https://dx.doi.org/10.4324/9781351314367-5), *Behavioralism in Political Science*, Routledge, pp. 93–108, ISBN 978-1-351-31436-7, retrieved 2020-12-14

8. Bauböck, Rainer, "Normative political theory and empirical research" (https://dx.doi.org/10.1017/cbo9780511801938.004), *Approaches and Methodologies in the Social Sciences*, Cambridge: Cambridge University Press, pp. 40–60, ISBN 978-0-511-80193-8, retrieved 2020-12-14
9. Leftwich, Adrian (2004). *What is politics?: the activity and its study*. Oxford: Polity.
10. Gray, JW (July 19, 2011). "The Is/Ought Gap: How Do We Get "Ought" from "Is?" " (https://ethicalrealism.wordpress.com/2011/07/19/the-isought-gap-how-do-we-get-ought-from-is/). *Ethical realism*. Retrieved December 14, 2020.
11. Shaver, Robert. “Hume's Moral Theory?” History of Philosophy Quarterly, vol. 12, no. 3, 1995, pp. 317–331., www.jstor.org/stable/27744669. Accessed 14 Dec. 2020.
12. Leigh Price (2019) Introduction to the special issue: normativity, Journal of Critical Realism, 18:3, 221-238 [1] (https://www.tandfonline.com/doi/abs/10.1080/14767430.2019.1638133)
13. Korsgaard, C. (1992). "The Sources of Normativity (https://tannerlectures.utah.edu/_documents/a-to-z/k/korsgaard94.pdf)" (PDF). *The Tanner Lectures on Human Value*.
14. *The Globalization of World Politics: An introduction to international relations*, New York, Oxford University Press ISBN 9780199297771, Fourth edition, pp.2-13

Further reading

- Canguilhem, Georges, *The Normal and the Pathological*, ISBN 0-942299-59-0.
- Andreas Dorschel, 'Is there any normative claim internal to stating facts?', in: *Communication & Cognition* XXI (1988), no. 1, pp. 5–16.