

Mapping Moral Pluralism in Behavioural Spillovers: A cross-disciplinary account of the multiple ways in which we engage in moral valuing

Michael Vincent

University of Queensland and Australian Catholic University, Australia

Ann-Kathrin Koessler

University of Osnabrück, Germany

Author Note

We thank Fabian Thomas, George Nguyen, Talia Sellers, Ben Retschlag, and William Sked for their valuable feedback on earlier versions of this manuscript. We are also very grateful for the suggestions made by several anonymous reviewers. Ann-Kathrin Koessler acknowledges the support of the Alexander von Humboldt-Foundation.

Corresponding Author: Michael Vincent, School of Historical and Philosophical Inquiry, University of Queensland, E-mail: m.vincent2@uq.edu.au.

Abstract

In this article, we reflect critically on how moral actions are categorised in some recent studies on moral spillovers. Based on classic concepts from moral philosophy, we present a framework to categorise moral actions. We argue that with a finer classification of the moral values, associated behaviour is better understood, and this understanding helps to identify the conditions under which moral licensing takes place. We illustrate our argument with examples from the literature on pro-environmental behaviours. Moral spillovers are frequently studied in this behavioural domain and to understand what causes their occurrence is highly (policy) relevant if we wish to promote sustainable behaviour.

Keywords: moral values; value pluralism; pro-environmental behaviour; behavioural spillover; moral licensing

1) INTRODUCTION: MORAL SPILLOVERS

Past good deeds can ‘liberate’ individuals to engage in behaviours which they otherwise would have considered to be inappropriate or immoral (Merritt, Effron, and Monin, 2010). For example, campaigns which have successfully promoted household water conservation have had the unintended effect of leading to increased household energy consumption (Jesso et al., 2017). This ‘moral licensing’ is an example of a ‘moral spillover effect’ much studied by behavioural scientists. There is still, however, great uncertainty about what causes a spillover from one im/moral action into future moral decisions.

In this paper, we argue that confusion about the forms that moral judgments take seems to be a key component in the failure to find agreement on what is occurring during purported cases of moral spillovers. This is particularly acute in cases like the aforementioned water conservation study, in which disagreement over moral spillover effects intersects with disagreement about the nature of environmental values. After giving a short introduction to the phenomenon of moral spillovers (Section 2), we illustrate how moral judgment can be modelled in a number of different ways (Section 3). Here we present models of moral assessment widely accepted in the philosophical literature, which we believe are plausible and utilizable by empirical researchers. Using examples from the recent spillovers literature, we illustrate how these concepts are already implicit in much empirical work, and how making them explicit can lend greater clarity (Section 4). This is expanded up in Section 5, where we hope to demonstrate some degree of irreducible complexity in moral decision-making: measuring morality cannot (always) be done with a simple spectrum from very bad to very good, and even apparently simple moral dilemmas may have more ‘contextual baggage’ than they first appear to. Ultimately, we discuss points which should be considered when eliciting and describing moral decision-making. Finally (Section 6), we present two models which can be adopted by researchers in order to clarify discussions within academic work and to improve the quality of data collected from human participants.

2) DEFINING MORAL SPILLOVERS

Spillovers are typically divided into *consistency effects*, in which an initial action begins a trend of similar action, or *balancing effects*, in which a good [bad] *behaviour 1* is followed by a bad [good] *behaviour 2* (Dolan and Galizzi, 2015 provide a good review, and we have imitated some of their terminology here; Truelove et al., 2014).¹ Our particular concern here is the study of ‘moral spillovers’, that is, spillovers in which moral belief or motivation seems to play a role (Nilsson, Bergquist, and Schultz, 2017 give a good taxonomy of spillovers, both moral and non-moral). Within this field, balancing effects are often referred to as cases of moral ‘licensing’ (a good act licenses a bad act) and moral ‘cleansing’ (a bad act is atoned for with a good act). *Prima facie*, consistency and balancing effects are incompatible with each other, so identifying the particular conditions under which one arises and the other does not is of great interest. This and similar problems are particularly pressing for those who study moral spillovers, as conflicting results and failures of replication have led to disagreement regarding the strength and prevalence of the phenomenon (Blanken et al., 2014; Blanken, van de Ven, and Zeelenberg, 2015). Sorting out the subtleties of moral decision-making is, we believe, an important step to determining what can now be said with confidence on the issue, and to making any plans for future replications or novel research.

We intend for the principles discussed to be applied to the study of all kinds of pro-social behaviours. However, in this paper we have focused specifically on studies in the domain of pro-environmental behaviour. Aside from being of particular relevance to our research interests, these studies have two advantages for our project: First, a fairly clear split between self- and other-regarding reasons-for-action exists in many pro-environmental actions. In at least some situations (such as recycling, or saving water while staying in a hotel), acting in an

¹ For the purposes of this paper, we focus on behavioural cross effects, i.e. spillovers to other behavioural domains. Behavioural effects within the same domain such as crowding effects are disregarded. For further readings on this matter we would like to direct the reader to e.g. Deci, Koestner and Ryan (1999), Frey and Jegen (2001); and for the environmental domain to Rode, Gómez-Baggethun, and Krause (2015).

environmentally friendly way has no obvious individual material benefits. (In others of course, such as reducing costly water and electricity use in the Jessoe et al., (2017) example, pro-environmental and self-interested behaviour can clearly coincide, and we may be able to explain much of the behaviour without appeal to any moral or other-regarding motivation.) So, we can identify more easily some cases which seem to clearly be cases of ‘moral’ behaviour, in the sense common to the literature of ‘pro-social’ ‘pro-environmental’, or ‘altruistic’ behaviour. Second, and related to the first point, there is enormous controversy about the nature of pro-environmental values (Schmidtz, 2015) – a controversy unlike that in other domains of moral behaviour. So, there is scope for researchers to be open-minded about what exactly the motivations and attitudes are of people engaging in pro-environmental moral behaviour, and advances in understanding spillovers here might help in forming a broader account of environmental values, by providing analogies and disanalogies with other domains of valuing.

What actually counts as a moral value is itself a fraught question, but we do not need to take a stance on the subtleties of the answer here. It will suffice to differentiate reasons of 1) narrow self-interest (typically material, status-seeking, or hedonistic) from 2) other-regarding (perhaps altruistic) preferences and 3) preferences for what is perceived as being just or righteous or intrinsically good, and to locate morality in the vicinity of 2) and 3). While pro-environmental behaviour may be motivated by type 1 reasons (say, if the price per gallon of water is so high that conserving water is the most cost-effective option available), our interest here is in pro-environmental behaviour of type 2 (say, conserving water for the good of others in one’s community) and type 3 (conserving water because wastefulness is seen as intrinsically vicious or immoral).

We are not the first to attempt to emphasize the role of moral values in interpreting spillovers (Cornelissen et al., 2013; Dolan and Galizzi, 2015; Jordan, Mullen, and Murnighan, 2011; Sachdeva, Iliev, and Medin, 2009), but we believe that the discussion could be improved if some ambiguities in the descriptions of moral values are cleared up. In the following, we

present ideas that have a long history of use in moral philosophy. However, the social science literature has not fully taken advantage of them yet. Acts or outcomes are regularly described as (morally) ‘good’ and ‘bad’ in the literature. From the point of view of the moral philosopher, such terms obscure important distinctions. The key lessons which we hope to demonstrate are first that humans value in many different ways, and there is an extensive, though fallible, vocabulary with which to discuss these moral values already (Section 3); second, this vocabulary should be a helpful complement or alternative to the current practices used for explaining the workings of spillover effects (Sections 4 and 5); thus, there is a ready program available for those who are concerned to further explore the alleged connections between moral valuing and behavioural spillover effects. It is important to note that our project is a conceptual contribution. We do not claim to have proven that our classifications of moral judgments is the best available for research on moral psychology– to prove this could require a whole academic sub-discipline worth of empirical work (and indeed, in Appendix 1 we offer a case for the incompleteness of this scheme). Nor do we seek to rely on the authority of any philosophical tradition to prove the usefulness of these ideas: we have been selective in our presentation, taking only those concepts that seem useful and breaking with philosophical and economic orthodoxies where appropriate. Our aim is only to demonstrate that there is conceptual confusion in the existing literature and that there are resources available to clear this confusion up.

3) WAYS OF VALUING

Humans always have and probably always will have different sorts of moral values. Consider the following analogy: Researchers are doing a study in physiology where subjects are ranked by size, but for some subjects ‘size’ means ‘weight’, for other subjects it means ‘height’, for others ‘shoe size’, and the differences between the sorts of measurements are not acknowledged. So, someone could be ranked 60 – but it is not clear whether that means 60 inches tall or 60 kilograms, etc. Similarly, it is not clear that the person with the biggest number

is the biggest person – both because someone who ranked 62 for height could still be *shorter* than someone who ranked 61 for weight, and because ‘bigness’ is ambiguous here – we are not sure how to commensurate the different measures into a single measure of who is ‘big’. Obviously, this research would be deeply problematic. But ‘morality’, we will try to show, is a notion at least as troublesome as that of ‘bigness’ here.

Our analysis is far from a complete account of the plurality of moral values. We, like Arias-Arevalo et al. (2018), and Lockwood (1999), seek to provide something of a taxonomy of value judgments, although unlike them we are concerned with people’s perceptions of pro-environmental *morality* rather than their perceptions of *the environment itself* (these are not, of course, totally unrelated subjects). Unlike them, we also place a stronger focus on distinguishing some common measures from each other in an operationalizable way, rather than providing a complete taxonomy. Hence, we have opted to present models which seem usable, rather than aiming for completeness (see Appendix 1 on some of the difficulties of achieving completeness in a taxonomy of moral judgments).² Lastly, we should differentiate our work from that of Katz-Gerro et al. (2017), who seek to identify *personality traits* which influence the *motivation to be moral*: our work might fruitfully interact with such studies, but we take no particular stance on that here.

a) How to Value: a Scalar model and a Categories model

‘Good’ versus ‘bad’ is a distinction which gives us some means for conceptualising spillover effects. But there are certainly other distinctions at play in ordinary moral judgments. First, we should acknowledge that some actions are taken to be not so much good, as ‘acceptable’ or ‘permitted’. Such actions contrast with those which are both bad and

² A reviewer has objected that in doing this, we have helped to perpetuate the ‘tacit assumption that there are only two plus one theories of morality’. (Perhaps as Joshua Greene is doing when classifying people into two ‘natural kinds’ (e.g. 2007: 37-9). This is a concern for us, and it is not our intention to promote such a simple account. We do suspect though that the popularity of the idea that there are only ‘two plus one’ accounts of ethics and our limited number of models here have a common cause: some hard limits on the ways on which we can elegantly model individual rational action.

‘unacceptable’ or ‘impermissible’. To say that an action is acceptable is not to suggest that it is a good one, but simply to say that it does not count as a bad one. Brian McElwee, in a recent paper, asks us to consider:

two ways of morally classifying acts: (a) a division into *deontic categories* of morally obligatory, morally forbidden, and morally optional (neither morally obligatory nor morally forbidden), and (b) an arrangement within an *evaluative scale*, ranging from morally best to morally worst. (2017: 505, *emphasis in original*)

It is fairly easy to visualise the latter scale, it might look like the two sided arrow displayed in Figure 1. We call this a ‘Scalar model’. [\[Figure 1 here – has been moved to end of doc in accordance with formatting guidelines\]](#)

The earlier categories, of obligatory, forbidden, and optional, can be represented as in Figure 2.1. or in Figure 2.2. We will call these ‘Categories models’. [\[Fig 2.1 and 2.2 here\]](#)

Where do the deontic categories (Figure 2.1 and 2.2) fit onto the evaluative scale (Figure 1)? It seems intuitive that the ‘forbidden’ acts will be towards the ‘worst’ end of the scale. But where do the ‘obligatory’ acts fit? Should they be represented around the middle of the spectrum? This area is also occupied by many acts which are not governed by obligations – like deciding which colour of t-shirt to wear today, or whether to eat an apple or an orange. Many of the ‘optional’ acts seem to be neither particularly good nor particularly bad. Working within common sense morality, we find that there is no neat way to graft the two ways of classifying acts onto one spectrum. It does not seem that one model is reducible to the other.

b) The Substance of these Models: Maximalist and Deontic Values, and Supererogatory Acts

One common way (it is not the only way that people employ) of assessing where an act falls on a scalar model of ethics is to evaluate its consequences – either the actual consequences, or those that could reasonably have been predicted. This usually relies on a ‘maximalist’ account of valuing. ‘Maximalist’ valuing is the sort of valuing which fits most neatly into an economic

account of human nature – for example, the figure *Homo economicus* is a maximiser, and so is the utilitarian.³ While there is no uncontroversial account of what humans do, can, or should aim to maximise, archetypal contenders include income, consumption, happiness or preference satisfaction. ‘Maximalist’ is broad enough to encompass both selfish valuing (I want to maximise my income) and altruistic valuing (I want to maximise global welfare).^{4,5}

Many goods that we pursue seem evidently maximalist. If I want money, then presumably the more money I can get the better it is for me. If I volunteer to help clean up litter on a beach and have a maximalist mindset, then I want to see the beach maximally clean. I will of course have to stop at some point, but any reduction in litter will translate fairly directly into greater success.⁶

However, this is not the only way in which we assess value, particularly with regard to moral action. Consider the two examples above, of money and litter. We may feel that the more money we can acquire, the better the outcome is for us. But we may also feel, at the same time, that it is particularly important that we get the money which we *deserve*: the money that we have worked for, or which we are entitled to for some other reason. Similarly, while I may regret having to leave the beach with some litter on it, I am likely to feel particularly bad about the situation if I am the one who put the litter there in the first place: if it is *my* litter which I am *responsible for*. Many judgments like these are captured by the label of ‘deontic’ valuing.

³ A reviewer has pointed out that a concern with consequences could take forms other than a maximalist mindset - one could for example have a *satisficing* mindset, or care only about *minimising harm*. This is quite correct, but we leave it to the reader to model these and other variations. It is not a threat to our thesis if the reader can conceptualise and observe even greater pluralism!

⁴ But the different sorts of ideals are only conceptually related, and may not even be compatible in practice. If person A wants to maximise global happiness and person B wants to maximise global suffering, then clearly these are incompatible maximalist goals.

⁵ The concept is also broad enough to encompass non-linear increases in value, such as cases of diminishing marginal returns.

⁶ It is worth noting that humans often make some strange and apparently inconsistent maximalist judgments, for example in the case of the ‘endowment effect’. See Knetsch (1994) on the importance of this in environmental values. See also Spash (2000) for a range relevant examples, some of which Spash suggests involve non-maximalist values being observed by researchers who set out to observe maximalist valuing.

The term ‘deontic’ suggests a sense of strict obligation, of moral *duty*. A duty implies a strong reason to engage in or refrain from certain acts. Sometimes this is presented as a limit on one’s set of options: ‘You shall not do x’ might suggest that a certain course of action should be unthinkable, or at least that we should be very averse to it, regardless of the potential consequences. It is common to suppose that the most important duties should be ‘negative duties’ or duties to refrain from certain forms of behaviour (‘thou-shalt-nots’); this is usually held to be because it is harder to demand consistent active behaviour from people – for example, while we can plausibly demand that a person live their entire life without killing another, it is harder to demand that they intervene at every opportunity to prevent another from dying.

‘Deontology’ has come to denote the discussion of a range of values which do not always provide strict obligations. While one feature of it is talk of duty and fundamental rights (and other notions which might be identified as ‘sacred values’ by psychologists (see for example Tetlock et al., 2000)),⁷ the term is also invoked for the discussion of rule-based moralities in general. As we saw above, not all moralising is concerned with maximising the good; rather, some is concerned with setting out the bounds of the acceptable. And it is here that rules (whether laws, or conventions, or social norms, or the edicts of moral leaders like Moses or Immanuel Kant) are helpful, because they reliably differentiate the acceptable from the unacceptable in a wide range of different circumstances. So in modern parlance, despite its etymology, ‘deontology’ has come to be seen as just as much the study of rules and rights as the study of moral absolutes.

To illustrate the differences here, let us briefly look at the example of donating to charity, an act which is frequently referred to as paradigmatically moral, and which is frequently the

⁷ Spash (2000) notes that some study participants, when asked to treat a matter of basic rights and duties as an economic transaction, made a ‘protest bid’ of no money at all, which could have been misinterpreted as an absence of concern. This looks very much like a case of ‘sacred values’, and Spash’s own description is akin to ours here.

measure of ‘good’ behaviour in the spillovers literature (e.g. Sachdeva, Iliev, and Medin, 2009). Donating can be seen as a maximalist act – other things equal, it is a good thing to donate, and the more donated the better. So, donating four per cent of one’s income is roughly twice as good as donating two per cent.⁸ We understand that this is generally the interpretation of the act used by spillovers researchers. Contrast this with the historical practice of paying a religious tithe (at least at times when this was not strictly enforced), or the contemporary practice of taking the ‘Giving What We Can’ pledge (Giving What We Can, n.d.). In both cases there is an expectation that ten per cent of one’s income will be given away, and there is a strong expectation that the ten per cent figure will be reached. While we can still talk of a ten per cent donation being better than a five per cent donation, there is also a very significant distinction between giving ten per cent and giving nine per cent. And this latter distinction is likely to be of a deontic sort, between the acceptable and the unacceptable, between success and failure in acting on one’s obligation.

With an understanding of the contrast between maximalist and deontic thinking, we can now introduce one final notion, that of ‘supererogatory acts’. Whereas duties are thought to set a minimum standard that we are expected to meet, and maximisable values set a target which is otherwise good for us to aim at, there is also a commonsense notion that some acts are not just morally acceptable but morally excellent. These are known as ‘supererogatory’ acts. To do one’s duty is usually not considered to be a supererogatory act. But duties help to set a standard for supererogatory acts – they are acts that go ‘well beyond the call of duty’. Conversely, a purely utilitarian, maximalist, account of morality cannot make much sense of the notion of ‘supererogation’ – achieving the best possible outcome is the whole point, and there is nothing better to do than that. We should distinguish two aspects of the notion of supererogation, both of which need to be recognised as features of commonsense morality. We have already stressed the excellence of some acts, but another feature of them is their *unnecessariness*. People are not

⁸ Donations from individuals in the US are generally around 2% of disposable income (Perry, 2013).

typically *blamed* for failing to perform supererogatory acts. The paradigmatic case of a supererogatory act is one which is both excellent and is not required by our day-to-day duties.

Supererogation in particular is a notion which seems to be wholly absent from the spillovers literature. To return to the example of donating 10% of one's income: to many people this will appear to be an act of generosity which they think is admirable but which they have no intention of imitating. This is a greater act than, for example, spending a few dollars more on 'green' product or donating a few dollars to charity. But still, for a person with a low disposable income, even a small amount of money can be of great use, and giving it up can be quite a sacrifice. In spillovers experiments, when giving is used as outcome measure, all three sorts of valuing can thus be present: There is a notion of deontic thinking as participants should only accept payments which they deserve according to the rules of the game (e.g. Mazar and Zhong, 2010). There is a notion of maximising good outcomes as participants who donate more money are taken to be acting in a better way than those who donate less (e.g. Sachdeva, Iliev, and Medin, 2009). There may also be participants who believe they are engaged in or are reflecting upon a supererogatory act - although researchers, in our experience, do not consider this possibility. Not only do individuals assess opportunities for prosocial or pro-environmental acts differently, they will likely import their own value assumptions into their interpretations of others' decisions. A simple implicit failure to agree on what sorts of values are in play can, we believe, distort the design and analysis of research studies: leading us to exaggerate the similarities between studies, to conflate important variables, or to misunderstand participants' self-reporting of their experiences and beliefs.

c) Domains of Action

There is a final issue of value pluralism which is worth making explicit, as its importance is not always recognised. This is the question of how to go about promoting one's values. People who value the environment might refrain from damaging it themselves, they might encourage others to refrain, or they might actively engage in environmental protection. These are already

three ways in which an individual can live up to upon her environmental values. But environmentalism might also affect one's diet, career choice, spending (whether consumable or philanthropic), voting and other civic behaviour, travel, even friendship networks or aesthetic preferences.

We will call this consideration the 'domains of action'; that is, the plurality of ways in which an individual acts on or expresses her values. In doing this we imitate the use of the term 'domain' in the spillovers literature. In the literature, domains differentiate certain regions of concern, such as environmentalism or honesty. So researchers might refer to the 'environmental domain' of moral behaviour, or the 'environmental domain' might be broken down into constituent domains such as the 'energy domain' and 'recycling domain'. For example, Noblet and McCoy (2017) sought to study decision making in the 'energy domain', while Mazar and Zhong (2010) did 'cross-domain' studies to examine the spillovers from the domain of environmentalism (green consumption) into the domain of honesty.

To illustrate the difference between the domain of interest and the domain of action let us use an example from the domain of interest 'environmentalism': the concern for recycling. People concerned (for ethical reasons) about recycling can respond to this interest in various domains of action, such as changing their consumption behaviour, campaigning for governance reform, or encouraging others to recycle more. And, depending on the individual's worldview, some of these actions may appear more attractive or appropriate. In experimental studies, however, subjects are often confronted with only one possible domain of action. Methodologically, this increases controllability and makes it easier to identify causal effects. However, if the motivations and values that the subjects associate with the particular action domain are not measured and taken into account, a spillover effect might be found that would have not occurred for a different action domain or for another population. Campaigning for government reforms, for example, may in a student sample generally perceived as something valuable, while this view might be less prevalent among the general population. On the other hand, it is also possible

that in reality spillover effects will arise in a completely different action domain than the one selected for the experiment. In Section 5.b, we will discuss, on the basis of existing spillovers studies, how study results may be compromised by not considering the full range of ‘domains of action’, that is, the ways in which people act on their values or motivations.

4) MORAL CONCEPTS IN THE CURRENT SPILLOVERS LITERATURE

Some of the above concepts appear in a few existing studies of behavioural spillovers. Cornelissen et al., (2013), for example, link a ‘rule-based mind-set’ to consistency spillovers (both good and bad). ‘Moral rules... do not naturally lend themselves to... trade-offs, because ‘a rule is a rule’ (Dolan and Galizzi, 2015: 9; summarising Cornelissen et. al.’s (2013) findings). This ‘rule-based mind-set’ is what we call deontic valuing (with elements of ‘aretaic valuing’, for details please see Appendix 1). Mullen and Monin provide an excellent summary of the deontic distinction between unacceptable and acceptable when they write that ‘perfect duties are black-and-white litmus tests, but they are asymmetrical: Someone violating a perfect duty is immoral, whereas someone respecting a perfect duty does not get much moral credit.’ (2016: 378) Similarly, Monin and Miller (2016) suggest a distinction between ‘moral opportunities’ and ‘moral tests’, the former being something like an opportunity to perform a supererogatory or maximalist act, the latter being a case of one’s being under an obligation to perform a duty.⁹ And finally, Truelove et al. (2014) talk of different ‘decision modes’, including ‘calculation-based’ (maximalist), ‘role-and rule- based’ (deontic, with an aretaic element).

In other research papers which focus on the general behavioural spillovers the distinctions in moral values we listed in Section 2 play no part. We should stress that this is not a problem for the work – values are controversial and can be difficult to pinpoint. But caution is needed when the results are subsequently interpreted and compared with other studies. Carrico, et. al. (2017), for example, discuss spillovers using a distinction between ‘morally good’ and ‘morally

⁹ That is, a ‘perfect duty’, to use Immanuel Kant’s term, a term which also appears in the previous quote from Mullen and Monin (2016).

dubious', as well as using the term 'pro-social behaviour' which is seemingly interchangeable with 'morally good', and the term 'pro-environmental behaviour' which is seemingly a subset of these. This language use overlooks the middle ground of 'moral acceptability' between the good and the dubious, and fails to distinguish moral excellence from more mundane forms of moral good. Further, the assumption that pro-environmental behaviour is always prosocial is questionable.¹⁰ Similarly, in other cases, the focus is on 'environmentally-friendly behaviour', with the link between this and morally correct behaviour being vague but unmistakably present. Lacasse (2017), for example, is interested in the link between 'Pro-Environmental Behaviours' and political attitudes, internal motivations, personal norms, and environmental identity, all of which seem intimately linked to subjects' sense of what is 'moral'. The paper provides an interesting examination of what we call 'domains of action', though in all other respects the different 'ways of valuing' we have described here are not discussed.

5) ROOM FOR IMPROVEMENT IN THE CURRENT USAGE

For other research studies, which focus more specifically on the moral dimension in behavioural spillovers, we believe that applying our value scheme can shed light on research gaps which currently exist. We can illustrate this with some cases of cross-domain spillovers, in which an initial behaviour affects the likelihood of an act in a very different 'domain', or area of moral concern.¹¹ Our examples here are primarily drawn from those who have followed-up on the work of Sachdeva, Iliev, and Medin (2009). Generally, our concerns are that:

- 1) Often, it is not clear how distinctions between moral, amoral, and immoral behaviour are drawn, and so whether the researchers participating in the dialogue themselves agree on the extension of these categories. (Section 5.a)

¹⁰ See Routley (1973) for a classic attempt to rebut this.

¹¹ Environmentalism perhaps being one moral 'domain', or a catch-all term for a number of domains like sustainability, litter reduction, species preservation, etc.

2) Researchers may interpret a subject's 'bad behaviour' as a (temporary, or induced) lack of interest in acting morally, when the behaviour in fact indicates different beliefs about what 'good behaviour' is. In other words, disagreement regarding which 'domain of action' is appropriate – what sort of action must be taken to do the right thing – can be misconstrued as apathy or immorality. (Section 5.b)

3) Researchers may talk of the severity of wrongdoing, while subjects may not agree with their severity assessments. So what appears to a researcher to be inconsistent behaviour ('moral balancing') may be consistent in the eyes of the subject. (Section 5.c)

To illustrate these concerns we will work through exemplary passages from this literature.

a) Delineating Im/morality

Consider this passage by Sachdeva, Iliev, and Medin, in which we are presented with distinctions between moral, amoral, and immoral, but not given enough guidance regarding where the authors believe that the distinctions are to be found:

The present research also does not address the question of whether people simply refrain from engaging in moral behavior in cases of moral licensing or are actually more liable to behave immorally. This is an important question to consider because the answer may indicate the strength of the licensing effect in motivating moral behavior. One possibility is that the licensing effect arises because of an accrual of "moral currency," which allows people to more or less passively engage in more secular sorts of activities until that currency has been spent. However, the more insidious possibility is that moral licensing lowers the bar of what is considered to be an amoral activity so that people are more likely to do immoral things that yield various types of secular benefits. These experiments show the first possibility to be true, because people refrained from doing something outwardly good (...), but our experiments do not show if moral licensing allows people to do something that is *prima facie* bad. A way to test this possibility might be to use cheating as a dependent variable and examine whether morally licensed individuals are more likely than others to cheat on a task that entails some sort of instrumental benefit. (2009: 527)

This is an important passage, as it helped to trigger a series of studies attempting to observe 'licensed' *immorality*. Following from Sachdeva, Iliev, and Medin's paper, Jordan, Mullen, and Murnighan (2011: 702) seek to find 'whether moral behavior can license actual immoral

behavior (rather than simply reducing the prevalence of moral behavior).' Jordan et al. are wary of a spectrum model of im/morality, writing that 'moral and immoral behaviors are not opposite ends of a single scale but rather two, distinct dimensions' (2011: 703). Note that all of these quoted sections seem to presuppose a 'categories' model of ethics, as in our Fig. 2.1 and 2.2. The examples given of immoral behaviour are acts like cheating and theft, in which the source of the 'immorality' does not seem to be a failure to maximise good outcomes, but rather the breach of an important rule of conduct. So, 'immoral' acts are identified via deontic reasoning. Whereas 'moral' acts included 'donating to charity, donating blood, volunteering' and 'help[ing] other people' (Jordan et al., 2011: 705) which we submit would typically be taken to be voluntary maximalist acts – on a categories model these are merely 'permitted' or 'optional', but they seem to fall somewhere on the 'good' side of a scalar model.

So when interpreting Sachdeva et al. and Jordan et al., we find a sort of 'middle ground' between rule-breaking (labelled immoral) and utilitarian maximalisation (labelled moral), which the writers label 'amoral' or 'nonmoral'.¹² 'Amoral/nonmoral' here could mean anything from behaviour which is unusually selfish but does not actually 'break the rules' to perfectly morally neutral behaviour (like choosing one flavour of icecream rather than another). And this ambiguity is not helpful if we are seeking to generalise about how spillover effects operate. It looks in both of the above cases as if maximalist and deontic thinking is being conflated, and we may be able to get some greater clarity if a distinction is introduced.¹³

¹² At another point Jordan et al. differentiate 'absence of moral behavior' from 'explicit immorality' (2011: 706), which in the context seemingly also identifies their amoral/immoral distinction.

¹³ Mazar and Zhong (2010) have also taken up Sachdeva, Iliev, and Medin's suggestion to 'use cheating as a dependent variable and examine whether morally licensed individuals are more likely than others to cheat' (Sachdeva, Iliev, and Medin, 2009: 527). They found (as did Jordan et al., 2011) that lying and cheating can be among the spillovers of some 'good' behaviours. In this study, Mazar and Zhong take some steps to assess their subjects' own interpretation of the link between ethics and pro-environmental behaviour (in Experiment I) and limit their own discussion of their findings to describing cheating as 'unethical' and 'socially undesirable'.

b) Domains of Action and Domains of Interest

We have seen that as well as dividing behaviour into domains by topic of moral concern (e.g. environmentalism, pro-sociality, personal integrity) we can recognise different interpretations of moral imperatives (maximalist, deontic, etc.). We should also recognise a distinction, between moral concern for actions, or for personal character, or for institutions, or for states of the world. It can be perfectly consistent to think that a certain act is wrong, even that the act is deontically prohibited, but also to believe that we should not interfere with or punish the bad act. A classic example might be the betrayal of a promise to one's spouse – it is the wrong thing to do, but that does not generally mean that outsiders have any business getting involved and making sure that the promise is kept. Furthermore, as Cristina Bicchieri (2006) has demonstrated, people can frequently disapprove of a social norm while continuing to participate in *and enforce* it – so judgment does not translate directly into action, and certainly not into collective action.

Similarly, we may disapprove of a state of the world or a pattern of behaviour, such as our disapproval of water or energy wastage, while also disapproving of some of the measures taken to fix it. Noblet and McCoy overlook this in their study on energy saving measures. They write:

[W]e focused on the potential for domain-specific licensing and hypothesized that individuals who participated in prior sustainable behavior *in the energy realm* would be less likely to support a future sustainable choice *in the energy realm* because they feel licensed by their prior behavior. (2017: 16, *emphasis in original*)

In this case, behaviour 1 is (remembering) the household's energy-saving behaviour, while behaviour 2 is to choose whether to support a government policy which would increase energy bills but would also increase renewable energy investment and/or fund energy efficiency programs. But voting for or advocating a certain policy is not part of the same decision mode as acting on related concerns in private. While these are two decisions within the 'energy' domain of interest, they are not within the same 'action domain' – these are two very different sorts of moral decisions about energy use.

Furthermore, the authors list as a limitation of their study the fact that ‘With our data we are unable to determine if participants followed through on their choice to support the Energy Policy Scenario by actually donating funds to renewable energy or energy efficiency.’ (2017: 17-8). Again the same reasoning applies. Just as being energy-conscious at home is not the same as supporting an efficiency-levy, supporting an efficiency-levy is not the same as donating money to an efficiency-charity.¹⁴ Imagine an anarchist environmentalist (or a less extreme form of ‘small-government environmentalist’). We ask them, ‘Will you buy some energy saving lightbulbs?’ They say ‘yes’, because they are an environmentalist. We then ask, ‘Will you support a government program which increases taxes to pay for more energy efficiency?’ They say ‘no’, because they are an anarchist. Their values here are not inconsistent, but the domain of action in which environmentalism is exercised is pivotal.¹⁵

Noblet and McCoy’s paper is concerned with ‘domain specific licensing’ in the ‘energy domain’ – but while examining this the study crosses back and forward between the ‘domains’ of household management, democratic citizenship, and charitable giving. Hence, what the authors observe does not need to be interpreted as a moral licensing effect, but rather shows that people seemingly behave inconsistently across different domains of action, perhaps because they are acting on multiple values.

c) Subjects’ Values versus ‘Objective’ Values

The final concern, which is tightly linked to the two previous points, is the fact that the value judgments of the researchers are often presented as correct and those of the spillover-prone subjects mistaken or biased. This is only one of many ways in which researchers’ values and moral views might shape their research design and analysis, but it is a particularly important

¹⁴ For an extended illustration of the difference between approving of a societal policy and acting on it unilaterally, and of approving of high taxes but not engaging in equivalent philanthropy, see Cohen (2000).

¹⁵ The work of Knetsch (1994) suggests another phenomenon that could be at play here: people recognize a significant difference between ‘willingness-to-pay’ for the environment to be protected and ‘willingness-to-accept’ compensation for its not being protected.

one to be aware of. Recall the previous quote in which Sachdeva, Iliev, & Medin (2009) suggested the possibility that ‘moral licensing lowers the bar of what is considered to be an amoral activity’ and so ‘allows people to do something that is *prima facie* bad.’ Now, it seems that ‘*prima facie*’ means from the perspective of the authors and their readers – it is clearly not from the perspective of the subjects, given the notion here of ‘lowering the bar of consideration’.

Commonly cheating and theft are judged to be particularly bad, probably because there are stronger norms in our societies governing theft or cheating than norms regarding giving to strangers or donating to charity. While refraining from these later actions is seen as ‘disappointing’, only theft is clearly *impermissible* (a *deontic* notion). It seems this is also the view of many (though not all) of the aforementioned researchers,¹⁶ given their tendency to single out the cheating as particularly ‘bad’ or ‘immoral’. However, it is an open question whether subjects would judge lying about the earned income in an experiment and cheating on a research fund as more immoral than not sharing given money with a stranger, or not donating to charity. Another example of the diverging interpretation of moral actions and its pitfalls are the aforementioned giving examples. While giving away a part of one’s earned experimental payoffs, either to a charity or to a fellow-subject, might be perceived by the researcher as a sensible way for the subjects to maximise good outcomes, a low-income subject (like an undergraduate student) might understand the act as a noble (‘supererogatory’) gesture which warrants some serious moral praise (or licenses some serious misconduct).

6) DISCUSSION AND CONCLUSION

Overall, the recognition that there are multiple ways in which humans engage in moral valuing suggests precautions for future research on behavioural spillovers.

¹⁶ See the examples given above in 5.a. Jordan et al. (2011) also use cheating as a paradigm example of an ‘immoral’ act.

Dolan & Galizzi (2015) have stressed the need for a way to quantify spillover effects, and have highlighted the lack of progress on this so far. We hope that we have demonstrated one important consideration of such a project: that measuring morality cannot (always) be done with a simple spectrum from very bad to very good. Rather, we should distinguish between the harmfulness and impermissibility, between actions and motives, and between acts, recollections, policy preferences, character traits, etc. We are aware that this is a very complex task, particularly since many factors are hard to objectively quantify, or at least to quantify in an intuitive or widely acceptable way. But we believe that also moral decision making is a highly complex process and future work considering this complexity can, in our opinion, help clarifying what leads to problematic behavioural spillovers.

To make a start, it should be helpful, when interpreting the work of others, to try to identify whether the work presupposes a scalar model of ethics (Figure 1) or a categories model (Figure 2.1 and 2.2.), or both, or some other schema. This can help to identify whether apparently similar research projects indeed arrive at comparable conclusions. Furthermore, when doing original research, researchers could ask subjects how they feel about certain sorts of im/moral behaviour with reference to the aforementioned models. For example, when discussing a certain act (such as cheating in a test or donating a certain sum to charity), subjects could be asked to fill out the following task sheet (note that here we have here added supererogatory acts to the categories model): [Figure 3 here]

We hypothesise that some topics will be more fruitfully discussed and classified with reference to one model than to the other. These will include our judgment, seemingly implicit in the work of others in the literature, that cheating and lying are better seen as ‘forbidden’ rather than as sitting somewhere on a scale of badness, and that while donating to a good cause is widely recognised as ‘good’ in the scalar model’s sense, discussion of its being obligatory, optional, or supererogatory will be more problematic (barring exceptional cases). We will be able to see this if subjects can independently agree on how to classify an act on only one model,

and the agreement is on a matter of substance. For example, everyone might agree that a \$1000 theft goes in the ‘forbidden’ category, but they might not be able to agree on whether it is better or worse than marital infidelity, lying under oath, or illegal polluting; so, they will not be able to agree on where these go on the ‘scalar model’. Conversely, while subjects may disagree about whether a donation of a given sum to a good charity is optional, excellent, or even obligatory, they should be able to agree that a larger donation is better than a smaller one (assuming that the task is well-framed by the researchers, and subjects do accept that the charity is indeed a good one). In this later case, we suspect, the scalar model will be more helpful in substantive assessments of just how invested in the ‘goodness’ of the good act the subject is, and so what degree of balancing or consistency might be occurring. If areas such as these of intersubjective agreement and disagreement can be found, then we will also have identified an important component needed in future work like that of Cornelissen et al. (2013), which seeks to encourage a rule-based (deontic) or consequentialist (maximalist) mindset in subjects: it will be easier to do so if the subjects are considering acts which are widely regarded as fitting into a certain category or are easily placed on a good-to-bad scale.

Another point that we hope to have brought to the reader’s attention is this: the assessment of spillover effects can only lead to generalisable conclusions if we researchers are more careful with the language that we use, so that we can describe exactly what we observe subjects are doing. Interpreting these observations then demands that the possibility of different perspectives must be taken into account. Part of this is explicitly differentiating between the values of researchers and the values of subjects. While it is important to observe that a subject’s evaluations are inconsistent, we should be sure that this is real inconsistency and not apparent inconsistency due to researchers’ values intruding.

Our theme has been the ‘dimensions of difference’ within morality. We believe that huge potential for cross-disciplinary work exists here. As Jordan, Mullen, and Murnighan suggest, the results of moral spillover experiments ‘have interesting implications for how people

cognitively categorize fairly disparate activities (e.g., helping and cheating) within a single moral rubric.’ (2011: 710). And, we hope to have provided with our framework how to categorise moral actions a helpful starting point for future collaboration which seek to understand pro-environmental behaviour and the determinants of potential moral spillovers.

FIGURES

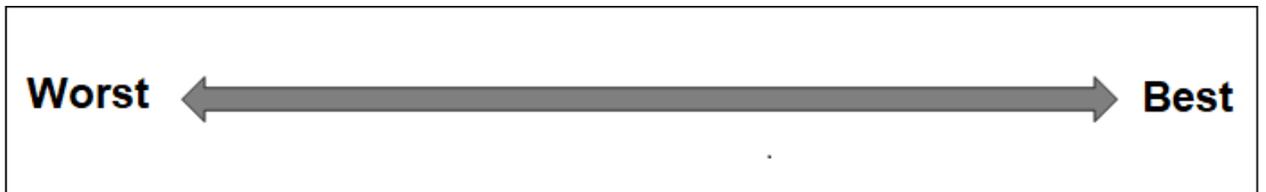


Figure 1: The Evaluative Scale

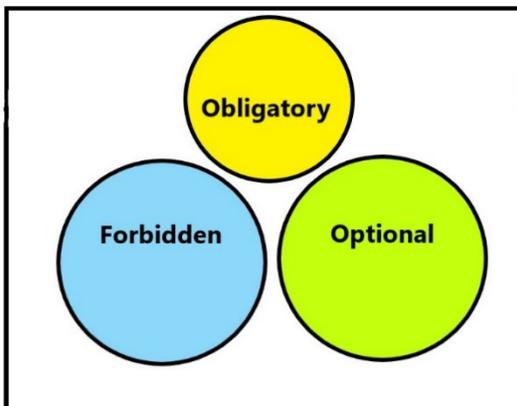


Figure 2.1: The Deontic Categories 1

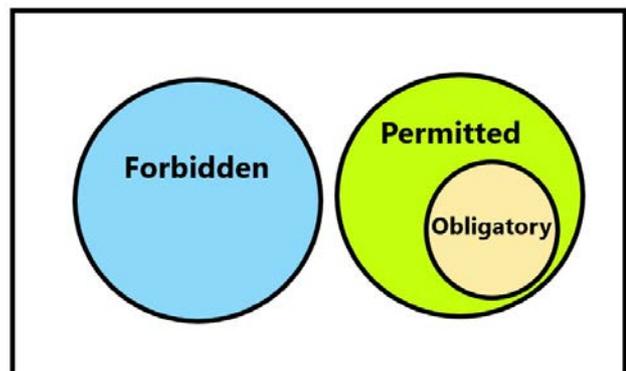


Figure 2.2: The Deontic Categories

2

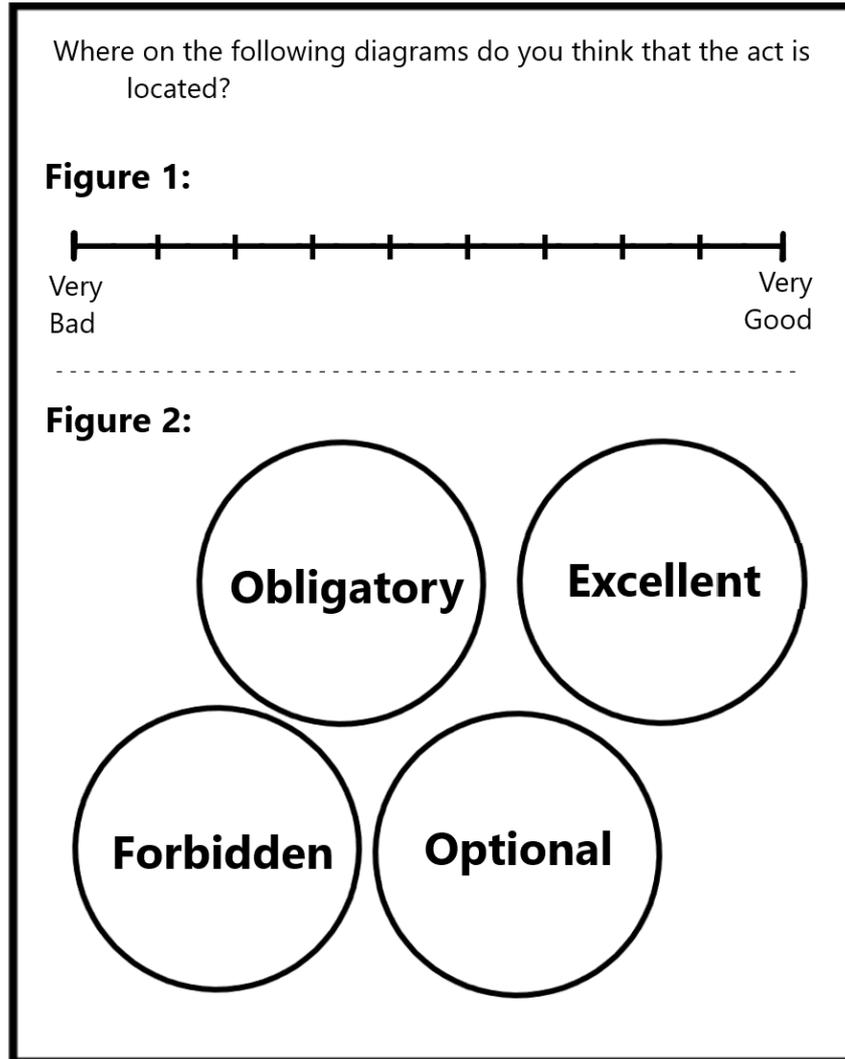


Figure 3: Task Sheet for Subjects¹⁷

¹⁷ Note that we have been careful not to label the centre of the scalar model: we suspect that many different judgments will be placed around there, and we do not want to leadingly suggest to subjects that the actions they judge to be 'amoral', 'acceptable', 'pragmatic', or 'cost/benefit-neutral' belong here or on any particular side of the centre.

BIBLIOGRAPHY

- Arias-Arévalo, P., E. Gómez-Baggethun, B. Martín-López, and M. Pérez-Rincón. 2018. 'Widening the Evaluative Space for Ecosystem Services: A Taxonomy of Plural Values and Valuation Methods' *Environmental Values* **27**(1): 29–53.
- Bicchieri, C. 2006. *The Grammar of Society: The Nature and Dynamics of Social Norms*. Cambridge, UK: Cambridge University Press.
- Blanken, I., N. van de Ven, and M. Zeelenberg. 2015. 'A Meta-Analytic Review of Moral Licensing' *Personality and Social Psychology Bulletin* **41**(4): 540–558.
- Blanken, I., N. van de Ven, M. Zeelenberg, and M.H.C. Meijers. 2014. 'Three Attempts to Replicate the Moral Licensing Effect' *Social Psychology* **45**(3): 232–238.
- Carrico, A.R., K.T. Raimi, H.B. Truelove, and B. Eby. 2017. 'Putting Your Money Where Your Mouth Is: An Experimental Test of Pro-Environmental Spillover From Reducing Meat Consumption to Monetary Donations' *Environment and Behavior* **50**(7): 723-748.
- Cohen, G.A. 2000. 'If You're an Egalitarian, How Come You're so Rich' *The Journal of Ethics* **4**(1–2): 1–26.
- Cornelissen, G., M.R. Bashshur, J.R., and M. Le Menestrel. 2013. 'Rules or Consequences? The Role of Ethical Mind-Sets in Moral Dynamics' *Psychological Science* **24**(4): 482–488.
- Deci, E.L., R. Koestner, and R.M. Ryan. 1999. 'A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation' *Psychological Bulletin* **125**(6): 627–668.
- Dolan, P., and M.M. Galizzi. 2015. 'Like Ripples on a Pond: Behavioral Spillovers and Their Implications for Research and Policy' *Journal of Economic Psychology* **47**: 1–16.

- Frey, B.S., and R. Jegen. 2001. 'Motivation Crowding Theory' *Journal of Economic Surveys* **15**(5): 589–611.
- Gilligan, C. 1982. *In a Different Voice: Psychological Theory and Women's Development*. Boston, MA: Harvard University Press.
- Giving What We Can. n.d. 'The Giving What We Can Pledge' <https://www.givingwhatwecan.org/pledge> (accessed 19 February 2018).
- Greene, J. 2007. 'The Secret Joke of Kant's Soul'. In W. Sinnott-Armstrong (ed.), *Moral Psychology*, pp.35–80. Cambridge, MA: MIT Press.
- Jessoe, K., G.E. Lade, F. Loge, and E. Spang. 2017. 'Spillovers from Behavioral Interventions: Experimental Evidence from Water and Energy Use' *Economics Working Papers: Department of Economics, Iowa State University*. E2e 033. https://lib.dr.iastate.edu/econ_workingpapers/52
- Jordan, J., E. Mullen, and J. Murnighan. 2011. 'Striving for the Moral Self: The Effects of Recalling Past Moral Actions on Future Moral Behavior' *Personality and Social Psychology Bulletin* **37**(5): 701–713.
- Katz-Gerro, T., I. Greenspan, F. Handy, and Hoon-Young Lee. 2017. 'The Relationship between Value Types and Environmental Behaviour in Four Countries: Universalism, Benevolence, Conformity and Biospheric Values Revisited' *Environmental Values* **26**(2): 223–249.
- Knetsch, J.L., 1994. 'Environmental Valuation: Some Problems of Wrong Questions and Misleading Answers' *Environmental Values* **3**(4): 351–368.
- Lacasse, K. 2017. 'Can't Hurt, Might Help: Examining the Spillover Effects From Purposefully Adopting a New Pro-Environmental Behavior' *Environment and Behavior* **51**(3): 259-287.

- Lockwood, M. 1999. 'Humans Valuing Nature: Synthesising Insights from Philosophy, Psychology and Economics' *Environmental Values* **8**(3): 381–401.
- MacIntyre, A. 1981. 'The Nature of the Virtues' *Hastings Center Report* **11**(2): 27–34.
- Mazar, N., and C.B. Zhong. 2010. 'Do Green Products Make Us Better People?' *Psychological Science* **21**(4): 494–498.
- McElwee, B. 2017. 'Supererogation across Normative Domains' *Australasian Journal of Philosophy* **95**(3): 505–516.
- Merritt, A.C., D.A. Effron, and B. Monin. 2010. 'Moral Self-Licensing: When Being Good Frees Us to Be Bad' *Social and Personality Psychology Compass* **4**(5): 344–357.
- Monin, B., and D.T. Miller. 2016. 'Moral Opportunities Versus Moral Tests'. In J.P. Forgas, L. Jussim, P.A.M. Van Lange (eds), *The Social Psychology of Morality*, pp.40-55. New York, NY: Psychology Press.
- Mullen, E., and B. Monin. 2016. 'Consistency versus Licensing Effects of Past Moral Behavior' *Annual Review of Psychology* **67**: 363–385.
- Nilsson, A., M. Bergquist, and W.P. Schultz. 2017. 'Spillover Effects in Environmental Behaviors, across Time and Context: A Review and Research Agenda' *Environmental Education Research* **23**(4): 573–589.
- Noblet, C.L., and S.K. McCoy. 2017. 'Does One Good Turn Deserve Another? Evidence of Domain-Specific Licensing in Energy Behavior' *Environment and Behavior* **50**(8): 839-863.
- Perry, S. 2013. 'The Stubborn 2% Giving Rate' *The Chronicle of Philanthropy*. <https://www.philanthropy.com/article/The-Stubborn-2-Giving-Rate/154691> (accessed 4 March 2019).

- Rode, J., E. Gómez-Baggethun, and T. Krause. 2015. 'Motivation Crowding by Economic Incentives in Conservation Policy: A Review of the Empirical Evidence' *Ecological Economics* **117**: 270–82.
- Routley, R. 1973. 'Is There a Need for a New, an Environmental Ethic' In *Proceedings of the XVth World Congress of Philosophy*, pp.205-210. Sofia: Sofia-Press.
- Sachdeva, S., R. Iliev, and D.L. Medin. 2009. 'Sinning Saints and Saintly Sinners: The Paradox of Moral Self-Regulation' *Psychological Science* **20**(4): 523–528.
- Schmidtz, D. 2015. 'Value in Nature'. In I. Hirose and J. Olson (eds) *The Oxford Handbook of Value Theory*, pp.381–399. New York, NY: Oxford University Press.
- Spash, C.L. 2000. 'Ethical Motives and Charitable Contributions in Contingent Valuation: Empirical Evidence from Social Psychology and Economics' *Environmental Values* **9**(4): 453–479.
- Tetlock, P.E., O.V. Kristel, S.B. Elson, M.C. Green, J.S. Lerner. 2000. 'The Psychology of the Unthinkable: Taboo Trade-Offs, Forbidden Base Rates, and Heretical Counterfactuals' *Journal of Personality and Social Psychology* **78**(5): 853–70.
- Truelove, H.B., A.R. Carrico, E.U. Weber, K.T. Raimi, M.P. Vandenbergh. 2014. 'Positive and Negative Spillover of Pro-Environmental Behavior: An Integrative Review and Theoretical Framework' *Global Environmental Change* **29**: 127–38.
- Williams, Bernard. 2006. *Ethics and the Limits of Philosophy*. London, UK: Routledge.

APPENDIX 1: RELATIONAL ETHICS, VIRTUES, AND (EVEN) GREATER HETERODOXY

While discussion of the importance of rules and outcomes is a mainstay of contemporary moral philosophy, these notions do not exhaust the range of concepts available to us. There are many others, some of which are associated with schools of ‘virtue ethics’, ‘care ethics’, anti-rationalist approaches, and others. Although it proved difficult to incorporate these ideas into the body of our paper, we want to present them here for readers to consider. We suspect that much of this sort of valuing is going to be particularly resistant to quantification, for reasons to be explained, and hence we have not offered a model of it akin to the ‘scalar’ and ‘categories’ models. But we do not want our contribution to be read as implicitly downplaying the importance of these ideas.

It seems possible to judge someone’s ‘moral character’, with such a judgment *not* reducible to only an assessment of the person’s past actions. The term ‘aretaic values’ is derived from an Ancient Greek term often translated as ‘excellence’, or more moralistically, as ‘virtue’. Aretaic judgments are judgments of character – of traits which are supposed to attach to people over the long term. It is commonplace, though not ubiquitous, to associate virtuous character with sociability, so a virtuous person is sympathetic, respectful, civic-minded, or whatever.¹⁸ It is also common to associate these judgments with deontic or maximalist expectations (‘they are a law-abiding person’; ‘they are an effective mentor’), though it would be a mistake to see an aretaic judgment as simply description of past or even predicted behaviour. A character trait may exist, presumably, without it is being publicly displayed, and it is not displaying the trait which principally makes someone virtuous, but possessing it. The clearest case of this is when someone

¹⁸ Some exceptions to this are the ‘militant’ virtues of someone like Homer’s Achilles, and the ‘monkish’ virtues of a religious ascetic. See MacIntyre (1981) for discussion of these and a good general introduction to ‘virtue ethics’.

has a certain character trait (say, generosity) but no opportunity to display it (because they have nothing to offer to others).

Some of the behavioural spillovers literature seems to recognise this distinction between one's perceived character and one's actual actions. For example, Miller and Effron (2010) stress the importance of one's self-identity (one's perceived character) in predicting moral licensing. Some researchers like Sachdeva, Iliiev, and Medin (2009) seem to have been able to manipulate the self-image of subjects, such that they can admire someone (or themselves) as a 'generous person' (or similar) without actually *doing* anything generous at all. As the subjects are not actually performing a relevant action, but rather just by reflecting on virtues of character, this looks like 'aretaic' valuing (perhaps of a seriously wrong-headed sort, in this particular case).

Furthermore, some would say that being a good person is not so much about knowing the rules and having the motivation to follow them, as about being able to wisely and sympathetically balance one's relationships with others. A generation ago, the alleged neglect of this topic was central to Carol Gilligan's (1982) 'care ethics' rejoinder to Lawrence Kohlberg's account of psychological moral development. Taking relationships seriously should have consequences for the use of toy dilemmas like trolley problems and dictator games in research. Once we accept that the concerns of 'care ethics' are an important part of our moral practice, then the worry looms that moral dilemmas in a lab may be a poor guide to actual conduct, given that so much of ethical behaviour is embedded in one's social context. The framing of such dilemmas will also be of huge importance. Consider, for example, the difference between introducing a task to a participant as 'a dictator game with a stranger' as opposed to 'a distribution task with a fellow-student'. The framing here may trigger very different feelings about one's role in the task, and one's relationship to other participants.

The simple distinctions used in the bulk of this paper have been between right and wrong, good and bad, or permissible and impermissible. But virtues and relationships are both narrower and more complex conceptions of im/moral phenomena. Some of these concepts may distort the boundaries of our original, simple distinctions. We might say that an act was ‘beneficent’, ‘courageous’, ‘harsh but fair’, ‘devious’, ‘exploitative’, or intended to ‘save face’, to give a few of the many possible examples. Following Williams (2006) it is common to call these ‘thick concepts’ (as opposed to ‘thin’ concepts such as ‘right’ or ‘bad’). How important thick concepts are in contemporary moral practice is debatable, but what seems clear is that they are not simple concepts to master – one has to be both a competent language user and a competent social actor to see exactly how the term is properly applied. They may be hard to analyse, and any simple dictionary definition is liable to be a poor guide to their use in practice. Even the relatively simple notion of ‘generosity’ provides an example: a ‘generous’ person is not simply someone who gives resources away; the label suggests something about the motivations of the person, and the relationship between, or status of, the giver and the recipients. A relevant example of the use of thick concepts in spillovers research is the lists of virtues in Sachdeva et al.’s (2009) experiments, and Blanken et al.’s (2014) attempts to replicate these with the list translated into Dutch. The ‘thickness’ of virtues like generosity is likely to make them resistant to quantification and observation, and perhaps even translation, particularly if we want our conclusions to be widely generalisable. Still, the topic is there for those prepared to engage with it.