Debate: Ideal Theory—A Reply to Valentini*

Holly Lawford-Smith

Philosophy, Research School of Social Sciences, Australian National University

IN her 'On the apparent paradox of ideal theory', Laura Valentini combines three supposedly plausible premises to derive the paradoxical result that ideal theory is both unable to, and indispensable for, guiding action. Her strategy is to undermine one of the three premises by arguing that there are good and bad kinds of ideal theory, and only the bad kinds are vulnerable to the strongest version of their opponents' attack. By undermining one of the three premises she releases ideal theorists from the paradox, which is as follows.

(1)   Any sound theory of justice is action-guiding.
(2)   Any sound theory of justice is ideal.
(3)   Any ideal theory fails to be action-guiding.

Here I shall respond to Valentini in two ways. First, I shall show that *both* (1) and (2) are false. The falsity of either is sufficient to release ideal theorists from the paradox. Second, I shall show that Valentini's response to (3) can be extended, because her bad kinds of ideal theory do not necessarily fail to guide action. The cumulative effect of my arguments further strengthens Valentini's support of ideal theory. I will deal with premises (1) and (2) in Sections I and II, and (3) in Section III.

I. NORMATIVE THEORY MUST BE ACTION-GUIDING

The first premise of Valentini's paradox asserts that any sound theory of justice must be action-guiding. She comments further that 'a capacity for guidance is widely considered a necessary attribute of any sound normative theory'.

I intend for what I say in this section to apply to both theories of justice and normative theories in general (and will proceed in discussing both), but I understand that some might escape the conclusions argued for here by stipulating that by normative theory they just mean theory that is action-guiding. If that is

their preferred usage of 'normative' then they should take my claim to be restricted to theories of justice. It would be absurd to try to claim that theories in general are necessarily action-guiding, given the long tradition in which theories have done explanatory, justificatory, and descriptive work. The thought motivating premise (1) must be, then, that there's something distinctive about theories of *justice* that entails that they must be action-guiding. In this section I will argue that there isn't.

Valentini excludes those who take a fact-insensitive approach to justice from the scope of her article. Such people will not find the paradox compelling, she says, because they reject premise (1) in advance. The idea must be something like: theories of justice are made up of principles, and on a fact-insensitive view like Cohen's both the principles and the comprehensive theories they are part of are not conditional upon facts, and thus not necessarily action-guiding.[1]

Notice that this depends heavily on what we consider 'action-guiding'; even fact-insensitive, wholly unconditional principles might still have the capacity to guide action, even if only indirectly (e.g. by providing a standard—more on this in Section III). It also depends on our assuming a necessary connection between fact-sensitivity and action-guidance, a connection which is at least not obvious.[2] Nonetheless, because Valentini excludes fact-insensitive views from the scope of her discussion, so too shall I.

What John Broome calls 'unowned oughts' certainly seem to belong to the normative domain if they belong anywhere. These include propositions such as 'life ought not to be so unfair'.[3] Arguably, there is also room in normative space for 'oughts' about the past, such as 'World War II ought to have not happened'. Realizing these 'oughts' might be infeasible or even impossible, yet they seem normative nonetheless. If that is true, then not everything in normative space must be action-guiding, which bears on premise (1). But one way to allow for these kinds of 'unowned oughts' without their threatening that premise is to say that they are statements which reflect principles (which in turn make up comprehensive theories). Then it can be the case that the general principles or theories are sensitive to facts or are action-guiding, without it being the case that

---

[1]Cohen 2003; 2008.

[2]I'm inclined to think that talking in terms of Cohen's fact-sensitive/fact-insensitive discussion actually just muddies the waters in the ideal/non-ideal theory debate. There are two important questions to ask in the region of facts and principles. The first, the one Cohen deals with, concerns the application of conditional moral principles. Do conditional principles always end in unconditional principles when we ask for explanation as to why we hold them? (See Cohen 2003, 2008; Pogge 2008). No matter our answer to that question, there still remains what seems for debates in ideal theory the much more important question, namely the meta-ethical question of what makes it true that principles (conditional or unconditional) apply. If we knew what the truthmakers were, we could commit to fact-sensitivity in the broadest meta-ethical sense, because the ultimate principles would be made true by their truthmakers. And if we thought that meta-ethical fact-sensitivity was sufficient to action-guidance, we could confirm premise (1) of Valentini's so-called paradox.

[3]Broome 2009, Ch. 2. Broome does not ontologically commit to unowned oughts, but rather comments on the semantics of ordinary language usage.

each discrete statement is. 'Life ought not to be so unfair' may plausibly rest on the principle 'one ought to act in such a way as to redress brute bad luck', which is itself an action-guiding principle.

This escape seems to work in some cases, but not in all. The explaining-away of seemingly unowned oughts in terms of action-guiding meta-principles won't work for example in cases where a loved one is killed in a natural disaster. We have a strong intuition there that life really ought not to be so unfair as to have let this be the case, but there doesn't seem to be a meta-principle that can cover that statement. Certainly not the principle of redress just considered, because there's nothing anyone can do to redress my bad luck in losing a loved one (short of superficial compensation).[4]

To give examples of such propositions is to try to elicit the intuition that guiding action is *not a necessary job* of normative theory. I want to support that intuition by outlining one very important philosophical task which does seem to vindicate such propositions. That task is understanding the nature of justice, or equality, or fairness, or freedom and so on. Investigation into the nature of things is a common task of philosophy in general.

We often appeal to far-fetched thought-experiments to tell us about the nature of things. For example, we were satisfied that knowledge was justified true belief until Gettier came along with cases where each of justification, truth and belief were present yet for which we did not have the intuition that there was knowledge.[5] Similarly, some of us were satisfied with a materialist analysis of the mind until Jackson gave us the Mary case, and Chalmers gave us the zombie case, eliciting a dualist intuition.[6]

On this understanding, at least part of the role of normative theory is to investigate those things essential to it. The ability to paint a picture of an ideal society may well tell us what justice or equality *is*, and counterexamples to the idea that those societies are the just or equal ones may demonstrate a fatal flaw in our understanding of the nature of those things. And this is no trivial matter; it's no use clamouring for a just society if no one has any idea how to recognize one when we have it, and we can't orient ourselves toward change for the better if we fail to know its direction. Understanding the nature of justice tells us something interesting and important about the world whether or not a just world

[4]It has been suggested to me that the idea of 'unowned oughts' is a hangover from times when more people believed in God, that is, when there was thought to be an entity capable of owning those oughts. The idea is that the proposition 'life ought not to be so unfair' is true only if there's someone who can make it the case that life is fairer, namely God. Or, the concept of an unowned ought might come out of an ambiguity in our language: perhaps I conflate the idea that there are states of affairs which ought to be the case with the idea that *it would be better if* certain states of affairs obtained. Despite these neat explanations, I find myself attached nonetheless to the idea that there are some genuine unowned oughts. But of course anyone who defines 'ought' or 'principles' or 'obligation' in such a way as to apply only to agents will avoid that problem and thus one of the motivations of this section.

[5]Gettier 1963.

[6]Jackson 1982. Chalmers 1996.

is a world accessible from our own. (This presupposes realism about, for instance, justice, but we might also paraphrase it for certain kinds of instrumentalisms, and so on.) Unowned oughts like 'life ought not to be so unfair' might give us a standard against which to assess actual circumstances, and might cumulatively allow us to deduce a relational standard which tells us which states of affairs are better and worse than others.

Valentini may respond to this by trying to argue that while investigation into the nature of things is indeed an important part of normative theory, it is still action-guiding. She might claim that something can be *indirectly* action-guiding if it contributes in some way to guiding action even if only much later down the line. It is reasonable to think that investigation of, for instance, justice in normative theory plays such a role, because getting clear about what justice is will allow us to appeal to it with greater force.[7] But if being indirectly action-guiding in this way is sufficient for being action-guiding, then a lot of people who define themselves by *disagreeing* with the proposition that 'a capacity for guidance is . . . a necessary attribute of any sound normative theory' will come out as agreeing with it. Valentini quotes Cohen as saying that political philosophy is about what we should *think* rather than what we should do; but what we should think will have *implications* (at least on the view I sketch in this section) for what we should do, and thus be indirectly action-guiding. By allowing 'indirect' action-guidance, we risk obscuring an important theoretical distinction.

Thus the response in this section to premise (1) of Valentini's paradox, namely 'any sound theory of justice is action-guiding', is to say that it is either trivial or false. If it just depends on how we parse 'action-guiding' so that even the most meta-ethically fact-insensitive statements, principles and theories can still be (e.g. indirectly) action-guiding, then it is trivial, because no theory will fail to satisfy it. If we parse 'action-guiding' in the standard way, as providing some practical guidance with regard to action, then the premise is false, because we have seen that one important job of normative theory,[8] as for philosophy in general, is telling us about the nature of things, such as justice. Theories of justice can be about the kinds of oughts which do not have to take feasibility constraints seriously, such as unowned oughts, or utopian oughts, and they can do important

[7]Here I should acknowledge that I'm assuming a moral realist view of justice in favour of a constructivist view. On this kind of view, justice is something in the world which our concepts are able to track. On the alternative view, justice as a concept is the solution to a practical problem, namely, that of the distribution of resources. (See the discussion in Korsgaard 2003). I'm not sure whether the difference between these two approaches matters for my argument in this section. On the moral realist view, we face the standard problem of epistemic access to moral truths. But on the constructivist view, we face the problem of intractable disagreement. In either case, there is a concept of justice (whatever justice truly is, or whatever the best solution to the problem we happen to be facing is), which can be explored independently of its application to human action. That is all I need to be able to say, I think.

[8]Zofia Stemplowska (2008) also argues that normative theory (the subset which she identifies as ideal theory) has other jobs to do than guide action. She argues that it also allows us to understand what our values require of us in the situations we find ourselves in (p. 22), because it helps us to see through potentially distorting constraints (p. 21), allows us to reflect on values which are complex and therefore nontransparent (p. 20), and issues conditional requirements (p. 19).

work in helping us to understand the nature of things better, which is intrinsically intellectually valuable, as well as instrumentally valuable in allowing clearer discussion when we do come to the task of guiding action.[9]

II. IDEAL THEORY AS NECESSARY

Premise (2) in the paradox asserted that any sound theory of justice must be ideal. For Valentini, a theory is ideal in a non-technical sense if it specifies an ideal world which we should aim for and ideal in a technical sense if designed under idealized assumptions. She focuses on theories that are ideal in *both* senses.[10]

Valentini asks: 'how could we formulate judgments about the justice and injustice of society, let alone promote institutional reform, without having an ideal of what a just society would look like?' The idea seems to be that formulating an ideal is *necessary*, for two reasons. The first is that it provides the standard by which we can assess non-ideal circumstances, and the second is that it tells us the *direction* in which we should push institutional reform.

This condition seems to presuppose a certain view about ideals that, I would argue, is false (although Valentini herself may not be committed to this view, for I take it that the paradox is supposed to capture the general assumptions of political philosophers). The view is that there is a single 'ideal', and once we fix it, we can know the distance between it and the actual world, and know further that all incremental movement toward it is a good thing. A visual image might help: imagine a set of concentric circles. The innermost circle is the actual world, and the outermost circle is the ideal world. Now imagine that for each of the circles in between, if we can get to them, we get ever closer to the ideal world.

The mistake in this picture is two-fold. Firstly, it ignores the fact that there may be *multiple* ideals, all of which fix a standard in a different place. There might be a world which is ideal with respect to justice but not to equality, and another

---

[9]At the recent Manchester Metropolitan Workshops in Political Theory, several people argued for a kind of dual-role view where ideal theory and non-ideal theory are not mutually exclusive methodologies of a theory but rather complementary components. Theories are complex: they might seek to guide action, they might seek to give an account of perfect justice, they might do either or both. If that is the right view to take of their relation, then premise (1) is false, because it claims that *any* sound theory of justice will be action-guiding. Some will, but some won't. (Defenders of that premise might reply with the argument that a theory *must* guide-action, but *may* in addition do other things. But then the debate is not any longer about whether theories should be action-guiding rather than ideal, but about whether theories must be action-guiding even when they are ideal). See Culp 2009; Child 2009; Zuolo 2009.

[10]There is a genuine question to be asked here about the relation between Valentini's (2009) understanding of 'idealization' and Onora O'Neill's (2000) well-known distinction between idealization and abstraction. O'Neill understands 'abstraction' as the bracketing of certain truths, and as 'idealization' reasoning from false premises. These correspond in a general way to Valentini's conclusion about good and bad kinds of ideal theory, but the fact that she employs the terminology of 'ideal' suggests she may be proposing a distinction within O'Neill's second category. If so, where O'Neill takes idealization to be genuinely troublesome, Valentini argues that it's only bad to idealize *subjects* of a theory, and okay to idealize background conditions which can be reintroduced at a later stage.

world that is ideal with respect to equality but not to justice; it might be that fundamental goods have to be traded off against one another so that there is no single 'ideal world'. Secondly, it ignores the fact that there can be 'better' worlds without there being a 'best' world.[11] On that view, given a set of five alternative worlds, we can say which of them is better (the most morally desirable) without having an idea of what an overall best world would look like.

If it is true that we can know better without knowing best, then it is false that ideal theory is *necessary* to formulating judgements about the real-world, or the direction of institutional reform. Likewise if there are multiple ideals, it is false that ideal theory (assuming a singular ideal) is necessary to either of those two tasks. Perhaps this is just a confusion in ordinary-language. Perhaps when philosophers talk about ideals and ideal theory they just mean 'ideal enough', or 'better than this world along some dimension of measurement'. Indeed, that kind of understanding does fit better with the kinds of proposals being made by ideal theorists (Rawls doesn't claim, after all, that a world governed by his two principles is the best of all possible worlds).

I have suggested that on one common understanding of 'ideal', ideal theory is *not* necessary to political theorizing. In what remains of the section, I want to argue that there is a way of understanding the thought-experiments of Rawls and Dworkin (identified by Valentini as good kinds of ideal theory) as *not* ideal at all. If that argument goes through, it is a further refutation of Premise (2), because both Rawls's and Dworkin's theories allow the formulation of judgements about justice, and specify a direction of institutional reform.[12]

Both Rawls and Dworkin can be understood as using a particular mechanism to obtain an idea of a just world. Both engage a hypothetical scenario: for Rawls, it is the original position with its veil of ignorance; for Dworkin, it is the desert island with its lack of prejudice and equality of bargaining power. These are thought-experiments, and they are common methodology in philosophy. Thought experiments are in most cases equivalent to counterfactuals.[13] For instance, while Rawls formulates the original position as a thought-experiment, we can easily reformulate it as a counterfactual: if citizens had been in a situation such that they lacked certain character traits (e.g. were not risk-loving, irrational, motivated by envy, and so on[14]), they would choose certain principles of justice to govern their political institutions, namely the difference principle and the equality principle.

We can do the same for Dworkin's desert island scenario: if individuals were to find themselves on a desert island, absent of prejudice, and so on, then they would choose a certain distribution of resources, and agree to a certain scheme of insuring against bad luck. The important point is that counterfactuals are true or

[11]See e.g.: Broome 1991; Sen 2006.
[12]As Valentini (2009) herself acknowledges, see her Section IV.
[13]Although there are some exceptions (e.g. rigid designators).
[14]Rawls 1971, p. 143; O'Neill 2000, p. 72.

false of states of affairs. Counterfactuals such as 'Rhiannon would have lied to Brian, had he asked her how much she had spent shopping that day' are *properties* of actual individuals. We need to be careful to separate the truth of counterfactual conditionals from the truth of their antecedents. 'If $p \;\Box\!\!\rightarrow q$' (following Lewis' notation for 'if it had been the case that p, then it would be the case that q'[15]) might be true; it might well be true that *if* we were not blinded by envy, risk-taking, and imperfect rationality, we'd choose certain principles of justice. But asserting the truth of the counterfactual conditional is not the same as asserting the truth of its antecedent, that is, saying that it *is* the case that we're not blinded by envy, not risk-takers, not imperfectly rational. Who would assert something so plainly false as true?

Consider an analogue. If orchestras had conducted their auditions blind (e.g. with those auditioning obscured from the view of those selecting) before 1970, there would have been a greater proportion of women in the world's best orchestras at that time. The antecedent of the counterfactual is false, because orchestras did not begin the practice of blind auditioning until fairly recently. But it seems remarkably likely that the conditional itself is true, because the proportion of women in orchestras has increased dramatically since the introduction of blind auditions.[16] And something seems to follow from the truth of that conditional, namely that blind auditioning is something we should be committed to. It's just the same with the counterfactuals in Rawls and Dworkin. If their counterfactual conditionals are true—if we had been in a position to choose principles of justice, if we had been on the desert island—then they have some normative force. If people would choose certain principles of justice were they not made blind and selfish by their psychological limitations, then perhaps those principles of justice should be in place.

What I'm suggesting is that far too much attention goes into attacking the antecedents of the counterfactuals, with objections like 'yes, but people aren't really like that, so nothing follows for real people!' That criticism is misguided because it focuses on the antecedent rather than the whole conditional. What's actually important is (a) whether the counterfactual is true, (b) what it recommends, normatively speaking, and (c) whether what it recommends is accessible. A much more relevant criticism would be that the ideals painted by those theories (ideal in Valentini's non-technical sense) are impossible to obtain from the real world.

To be perfectly explicit on what is being suggested here: Valentini's technical definition of ideal was 'designed under idealized assumptions'. I am arguing that there are no idealized assumptions in Rawls or Dworkin, there are just counterfactual statements that are either true or false of individuals *as they actually are*. If the counterfactuals are true, a certain world is justified as a good one (i.e. is ideal in the non-technical sense). The important relation is between the actual world and the world justified as a good world by the truth of the

---

[15]Lewis 1973.
[16]Gladwell 2009.

counterfactual, and that is a very different relation to that between the actual world and the world where the antecedent of the counterfactual is true. We should be asking whether we can get to the good world justified by the truth of the counterfactual, not whether our world is or could be like the world where the antecedent of the counterfactual is true. The latter is irrelevant.

I should pause for a moment to comment on the scope of this claim. Certainly it is not the case that any ideal theory can be re-described in the counterfactual terms I have employed here. The fact that Rawls's and Dworkin's theories can be understood in that way is an artifact of the kind of idealizing they do, namely their counterfactual analyses of justice which use dispositional properties of persons. I expect that the argument here can be applied to any ideal theory of that kind. But there are other kinds of ideal theory: there's ideal theory which simplifies, by focusing only on certain truths (like the frictionless plane thought-experiment in the sciences), and there's ideal theory which tries to say what we should do, or what the world should be like, if we want to achieve some goal or other. For instance, if we want to reduce our carbon emissions 90% by 2050 then almost everyone will have to change their lifestyle, but we can generally expect not everyone to change their lifestyle in the desired way. (People often comment that strict compliance is ideal in the pejorative sense of impossible. I think it's more accurate to say that it's not worth what it would cost. Presumably it would require coercive incentive structures, an oppressive state, and so on).

To summarize, premise 2 maintained that 'any sound theory of justice is ideal'. In this section I have argued that two paradigmatic theories of justice, Rawls's and Dworkin's, are not ideal. They are ideal in the non-technical sense in that they recommend a world, but they are not ideal in the technical sense because they are not designed under idealized assumptions. They do not assert the truth of the antecedent of the counterfactual; they merely assert the truth of the counterfactual conditional. The conclusion is that we must either say that neither Rawls nor Dworkin give any sound theory of justice (which is implausible), or that it's false that a sound theory of justice has to be ideal in both the technical and the non-technical senses Valentini endorsed.

III. IDEAL THEORY FAILS TO BE ACTION-GUIDING

Premise (3) is the condition Valentini chooses to attack in her article. The premise voices the worry that theory idealized too much from real-world conditions will fail to specify action for the real-world, instead prescribing action only appropriate for an ideal world we do not inhabit. She summarizes the worry as follows:

> Principles designed assuming ideal conditions, so the worry goes, are ill-suited to guide action in the real world, where such conditions do not hold. Given that a capacity for guidance is widely considered a necessary attribute of any sound normative theory, the "guidance critique" seems to pose a serious threat to liberal egalitarian thinking as a whole.

Now it is worth restating that Valentini and I do not have different objectives in mind. Her goal was to rescue ideal theory from its critics by showing that *not all* ideal theory is susceptible to the criticism outlined in (3). I agree with her on that point, but I think the arguments in favour of that conclusion are stronger than she has allowed for. I have already argued that there is no paradox for ideal theorists, because both premises (1) and (2) are false. In this section I want to argue that the 'bad' kind of ideal theory she identified is not as a matter of necessity vulnerable to her argument.

Recall a parallel Valentini drew with subject matter outside normative theory, to support the distinction she made between good and bad kinds of ideal theory. She said that ideal theory is bad if it idealizes its *subjects*, which she suggested Kymlicka's theory of minority rights, Rawls's theory of international distributive justice, and economic theory assuming perfect rationality, all do. Ideal theory is good if it idealizes only in a way that allows complications to re-enter the picture at the level of application, as she suggested Rawls's and Dworkin's theories of domestic justice do, and as Galileo's 'frictionless plane' thought-experiment in physics does.

I think Galileo's thought-experiment idealizes its subject and yet is none the worse for it. It assumes away friction, then makes a prediction about how bodies should behave. When we apply the theory, we build real-world complications like friction back into the story. One way of understanding what we are doing is that we specify an ideal, the way things behave in simplified or ideal conditions, such as in worlds without gravity, and then we *relax* the assumption of ideality by reintroducing real complications, like friction. This seems to apply *equally* to Rawls's ideas about international justice, and economists' assumptions about perfectly rational individuals. Rawls imagines that societies are 'reasonably well-ordered', and then says how they should act toward each other, given that assumption.[17] Economists imagine that people are perfectly rational, and say what they should do, given that assumption.

While these theories specify ideal conditions and the requirements for action that follow from them, they are not divorced from reality in a way that makes them inapplicable. We can relax the assumption of ideality and so too the requirements for action. We can do that by estimating the distance between the ideal and the real, between the perfectly rational agent and the actual agent, the perfectly well-ordered society and the actual society. Given a basic understanding of what this distance is, the stringency of the requirements for what we should *do* will relax in a way appropriate to actual circumstances.

Valentini denies this, arguing that for Kymlicka, one has to simply start from scratch when considering non-ideal circumstances; his theory tells us what to do only *if* we ever find ourselves in the conditions specified by the ideal. But why not take the 'ideal' as a standard—nations being relatively well-ordered, and

---

[17]Valentini 2009, Sections V and VI.

behaving in such and so a way toward one another. Then (so long as approximation is theoretically unproblematic, which is controversial) we can use that standard to make judgements about non-ideal circumstances, and decide on issues of culpability.

This idea that we can relax ideal standards in so-called 'bad' kinds of ideal theory does not hold for all such theories. It looks like there will be at least *some* cases of ideal theory in which we are not able to relax the ideal, or the requirements that follow from it. To return to the economists' assumption of perfect rationality, imagine that an ideally rational agent can perform infinitely many steps in a given system of logic. Now imagine that an actual agent can perform arbitrarily many finite steps. It might be the case that one actual agent can perform more steps than another, and so is in that sense closer to the ideal; but in another sense both agents have finite capacities, which means they are both infinitely far away from the ideal, and cannot ever reach it.[18] I certainly concede this point, but argue that it is still possible to maintain that *some* of the instances Valentini discusses have ideals which can be relaxed, in which case the category of 'bad' ideal theory will be narrower than one might have thought. Some also might object that any relaxing of ideal standards is open to the problem of the second best.[19] This objection seems to have the most force in cases with non-linear values. For instance, values like equality between persons and human health, seem straightforwardly linear (the more people that are treated as equals, the better, the less health problems in a nation, the better) and so less open to worries about non-obvious second-bests. It gets trickier with values like freedom, because we want as much guaranteed freedom as possible, which means trading-off against state protection. But the fact that there's more than one way of relaxing an ideal standard doesn't imply that the standard can't be relaxed—rather it implies the opposite.

In Section II, I argued that the much more important question for ideal theory is not whether there is an abstraction of the subject or environment of a theory, but whether what justifies the non-technical ideal world is reasonable, or sound, and whether the ideal is accessible from our world. This is to argue to some extent also against O'Neill's distinction between abstraction and idealization; it doesn't matter if the assumptions are literally false (e.g. to assert the absence of predicates that do obtain, or to assert the presence of certain predicates that do not[20]). Furthermore, the distance between abstraction and idealization seems flimsy at best—we can re-describe bracketing as asserting the absence of certain predicates. Valentini tries to get around that problem by saying that the bad kinds of idealizations require falsehoods such that the truth cannot be reintroduced (they're part of the theory construction). But even in that case, it's open to a critic

---

[18]I owe this example to Jens-Christian Bjerring, and am grateful to him for helpful comments on this section of the paper.

[19]See e.g. Goodin 1995.

[20]O'Neill 2000, p. 68.

to argue—as I have tried to do in this section—that a theory built on falsehoods might specify a standard that we can nonetheless relax to account for the falsehoods. For example, we assume perfect rationality to figure out how agents should act in order to maximise their expected subjective utility, but we know that real agents are not perfectly rational, so we don't hold them to that standard but rather to a sort of relaxed version of that standard.[21]

IV. CONCLUSION

In this article I have argued that Valentini's attempt to rescue ideal theory from its critics can be strengthened. I argued that no premise of her paradox is true, which releases ideal theorists from the paradox in a straightforward way. Premise (1) fails because normative theory has an important job to do in telling us about the nature of things important to discussion in the normative space, a job which has little to do with direct action-guidance. Premise (2) fails because typical discussions of 'ideal' leave no room for multiple ideals or relative standards, and more significantly because some paradigmatic sound theories of justice are not ideal. Premise (3) fails not only for the reasons she gives, but furthermore because there is a way to argue that even the 'bad' ideals she mentions can sometimes provide guidance about action, by fixing a standard which can be relaxed to account for non-ideal conditions.

REFERENCES

Broome, John. 1991. *Weighing Goods*. Oxford: Blackwell.
Broome, John. 2009. *How to be Rational*. Mimeo., Oxford University (draft manuscript dated March 2009).
Chalmers, David. 1996. *The Conscious Mind*. Oxford: Oxford University Press.
Child, Richard. 2009. The dual-component model of justice. Paper presented at the Manchester Metropolitan Workshops in Political Theory, 2–4 September, 2009.
Cohen, G. A. 2003. Facts and principles. *Philosophy & Public Affairs*, 31, 211–45.
Cohen, G. A. 2008. *Rescuing Justice and Equality*. Cambridge, MA: Harvard University Press.
Culp, Julian. 2009. Two conceptions of a complementary relation between ideal and nonideal theory. Paper presented at the Manchester Metropolitan Workshops in Political Theory, 2–4 September, 2009.
Gettier, Edmund. 1963. Is Justified True Belief Knowledge? *Analysis*, 23, 121–3.
Gladwell, Malcolm. 2009. *Outliers*. Boston: Little, Brown.
Goodin, Robert. 1995. Political ideals and political practice. *British Journal of Political Science*, 25, 37–56.
Jackson, Frank. 1982. Epiphenomenal Qualia. *Philosophical Quarterly*, 32, pp. 127–36.

[21]This claim that ideal standards can be relaxed relies to some extent upon there being a bridge from ideal to non-ideal theory, and there seems to be only one proposal on the table about how that bridge might work. See (Murphy 2000). I don't have in mind a formal algorithm for the relaxation from ideal standards, so if that is necessary then my proposal will be vulnerable until a plausible bridge is established.

Korsgaard, Christine. 2003. Realism and constructivism in twentieth-century moral philosophy. Pp. 99–122 in *Philosophy in America at the Turn of the Century: APA Centennial Supplement to Journal of Philosophical Research*. Charlottesville, VA: Philosophy Documentation Center. Available at ⟨www.pdcnet.org/pdf/8Korsgaard.pdf⟩.

Lewis, David. 1973. *Counterfactuals*. Oxford: Blackwell.

Murphy, Liam. 2000. *Moral Demands in Nonideal Theory*. Oxford: Oxford University Press.

Stemplowska, Zofia. 2008. What's ideal about ideal theory. *Social Theory and Practice*, 34, 319–40.

Valentini, Laura. 2009. On the apparent paradox of ideal theory. *Journal of Political Philosophy*, 17, 332–55.

Rawls, John. 1971. *A Theory of Justice*. Cambridge, MA: Harvard University Press.

O'Neill, Onora. 2000. *Bounds of Justice*. Cambridge: Cambridge University Press.

Pogge, Thomas. 2008. Cohen to the Rescue! *Ratio*, 21, 454–75.

Sen, Amartya. 2006. What Do We Want from a Theory of Justice? *Journal of Philosophy*, 103, 215–38.

Zuolo, Federico. 2009. Bridging the gap between ideal and nonideal theory: the concept of self-efficacy in normative political theories. Paper presented at the Manchester Metropolitan Workshops in Political Theory, 2–4 September 2009.