

# On Metaphysical Cases against Political Theories

*'[...] the optimistic view – which seems to be at the heart of much metaphysical rationalism – [holds] that all good things must be compatible, and that therefore freedom, order, knowledge, happiness, a closed future [...] must be at least compatible, and perhaps even entail one another in a systematic fashion. But this proposition is not self-evidently true, if only on empirical grounds. Indeed, it is perhaps one of the least plausible beliefs ever entertained by profound and influential thinkers.'*

Isaiah Berlin (2000: 117-118)

## 1. Introduction

It has become increasingly popular to criticize political theories from the perspective of the metaphysical debate on free will. Smilansky (1997) was one of the first to approach egalitarianism from this angle, and argued that the viability of different egalitarian theories is dependent on our solution to the free will problem, that is, the kinds of egalitarianism that are admissible depend strongly on the conception of free will they presuppose.

Following up on this endeavor with a more specific claim, Van der Deijl (2013) has used metaphysical arguments to criticize luck egalitarianism. Luck egalitarianism and its variations, a popular group of positions developed among others by Dworkin (2002), suggest that individuals should be compensated for negative effects due to those factors for which they have no responsibility. This compensation should be paid for at the expense of positive effects for which they were not responsible. Luck egalitarianism therefore requires a notion of responsibility that can separate factors for which we are responsible from factors for which we are not: it must separate luck from merit. Hard determinists cannot offer such an

account, because they see actions as necessarily following from physical states; compatibilists have tried, but according to Van der Deijl to no avail. An adequate notion of responsibility can be held by libertarians, but that position is unpopular because it requires the rejection of determinism. Van der Deijl concludes that, unless libertarianism can save the day after all, luck egalitarianism cannot distinguish luck from merit. In that case luck egalitarianism would collapse into outcome egalitarianism.

Supporting the arguments of Smilansky and Van der Deijl is the silent premise that metaphysical theory takes precedence over political theory, in the sense that a political theory should presuppose a plausible metaphysical theory. Kang has called this silent premise the 'Posteriority Thesis' (2003).

The Posteriority Thesis (PT) is intuitively appealing: it seems obvious that a political theory should not incorporate entities that cannot exist metaphysically. The consequences of accepting this thesis are, however, not nearly as intuitive since acceptance of PT forces us to dismiss notions that have not only proven to be useful in political philosophy, but that also seem to be morally significant. It also risks paralyzing political philosophy until certain metaphysical issues have been resolved. Although none of these points are in themselves sufficient reason to discard PT, they are reason enough to inquire into its foundation.

This paper subjects PT to a critical evaluation. Its main aim is to show that the truth of PT is often assumed to be trivial, while it has important consequences and while there are alternative approaches to political theory that can do without it. I will proceed by assessing whether it is correct to dismiss a political theory on the grounds of not being consistent with

plausible metaphysical theories about free will, as Smilansky and Van der Deijl have done. It should be added that it is explicitly not my intention to confirm or reject PT, nor to defend or reject any of the political theories discussed by Van der Deijl or Smilansky on non-metaphysical grounds. I merely wish to show that the influence of metaphysics in political theory is not necessarily as decisive as is sometimes thought.

The next section will introduce PT. The argument will then consist of three parts. Section 2 shows that PT is usually assumed rather than justified, and argues that PT has undesirable consequences. Section 3 shows that it is possible to do political philosophy without adopting PT, using the example of John Rawls. This example illustrates that metaphysical theories carry less, if any force for political theory if one takes moral beliefs as one's starting point.

## 2. Introducing the Posteriority Thesis

This section aims to clarify the notion of the Posteriority Thesis, and explains why it is important to subject it to scrutiny. The term 'Posteriority Thesis' was coined by Sung-Hak Kang (2003) in response to Smilansky (1997). Smilansky discussed egalitarianism as defended by Cohen. Cohen had argued that he could introduce his views on political theory without trespassing into the notoriously complicated free will debate. These are some of Smilansky's arguments:

'(i) Firstly, we cannot accept the assumption that people obviously and uncontentiously have the type of free will required by such an egalitarian conception as Cohen's. (ii) Secondly, the egalitarian case has much to gain from the clarification of the importance of the free will problem for questions of distributive justice (and more specifically, from the recognition of the implications of the lack of libertarian free will). However, one's judgment about the possible widespread attractiveness of egalitarianism would be very much influenced by one's position on the 'Compatibility Question' of the free will problem. (iii) Thirdly, the *form* of egalitarianism which will be considered adequate crucially depends on the position taken with regard to the free will problem.' (Smilansky, 1997: 159-160)

In his reaction, Kang uncovers the Posteriority Thesis as an implicit element of Smilansky's argument. While Kang admits that congruence between metaphysical and political theory is a theoretical virtue, he asserts that there is no 'metaphilosophical mandate' (2003: 108) that dictates (a) the necessity of congruence between the two, or (b) that metaphysics takes priority in achieving congruence.

Now let us try to define PT. It is important to notice that PT goes further than asserting the following:

1. If a political theory presupposes a metaphysical theory, it must be a plausible metaphysical theory.

PT claims more than (1) because Smilansky says that political theorists who *do not* supplement their egalitarianism with a metaphysical addendum on free will *should* do so: the 'adequate form of egalitarianism' after all *crucially depends* on one's position in the free will debate (see citation above). In other words, the possibility of putting metaphysical debates aside when doing political theory is rejected. Our working definition of PT should also incorporate that Smilansky assumes a hierarchy: our political theory should be adapted to fit our metaphysics, not the other way around. A better definition of PT would therefore be the following:

2. The only viable political theories are those that are grounded in a plausible metaphysical theory.

Statement (2) will serve as our definition of PT.

An important corollary of this definition is that a violation of PT is enough reason to dismiss a political theory. Van der Deijl exemplifies this in his argument against luck egalitarianism. To show this, let me first represent his argument.

According to Van der Deijl, luck egalitarianism requires 'substantive responsibility' in order to distinguish luck from merit. Substantive responsibility, a term borrowed from Scanlon (1999), refers to the kind of responsibility we carry for autonomous choices (that is, choices that are properly our own, rather than ascribed to us by others).<sup>1</sup> It can be contrasted with attributive responsibility, which regards choices that others attribute to us but are not autonomous.

Van der Deijl then asks which of the three positions in the metaphysical debate on free will entails substantive responsibility. Hard determinism, which holds that determinism is true and that, as a consequence, free will does not exist, does not entail substantive responsibility because hard determinism does not accommodate for individuals as autonomous agents: all human actions are the necessary consequence of the preceding physical states. The same problem holds for compatibilism. Compatibilists hold that some kind of free will can coexist with the truth of the determinist thesis. Since the problem for hard determinists was that determinism and substantive responsibility are not compatible, the same problem holds for compatibilists. Libertarianism, lastly, seems to do better with respect to substantive responsibility. Libertarianism rejects the determinist thesis and holds that free will exists, thereby allowing us to see some choices as autonomous. Libertarianism can, therefore, be compatible with substantive responsibility. The problem with libertarianism is, however, that it is considered an implausible position to begin with, since determinism carries too much appeal to be rejected.

Van der Deijl's conclusion is therefore that either libertarianism is true, or substantive responsibility cannot be saved. The latter possibility would imply that hard determinists and compatibilists cannot distinguish between luck and merit, and luck egalitarianism collapses into outcome egalitarianism.

I submit that Van der Deijl subscribes to PT defined as (2). To qualify, Van der Deijl must hold that a political theory is not viable if it is not supported by a plausible metaphysical theory. This seems to be the case when he says that luck egalitarianism must collapse into outcome egalitarianism if one's metaphysical theory does not offer us tools to separate luck from merit, as is the case for hard determinists and compatibilists. That compatibilism and hard determinism cannot metaphysically ground substantive responsibility should be a sufficient reason for dropping luck egalitarianism, unless one is a libertarian. By putting the argument this way, Van der Deijl rejects the option of justifying substantive responsibility on non-metaphysical grounds. This is evident in his response to Knight. Knight (2006) had suggested that one might not need a metaphysical basis to support one's political theory; moral intuitions could perhaps suffice.

Van der Deijl answers as follows:

'However, the implicit premise underlying [Knight's] argument is surely false: if our intuition goes against a metaphysical conclusion, the metaphysical conclusion is to be rejected. The point of drawing on arguments from the free will debate for the distributive justice discussion is exactly that it may show us that some of our intuitions may be wrong.'  
(Van der Deijl, 2013: 28)

Because luck egalitarianism cannot be grounded in hard determinism or compatibilism, proponents of those metaphysical views cannot accept luck egalitarianism. If moral intuitions disagree with our metaphysical conclusions, the intuition should be rejected. Both show that Van der Deijl endorses PT: the only viable political theories are those that can be grounded in a plausible metaphysical theory.

### 3. Unwanted consequences of accepting the Posteriority Thesis

Recall that, according to Kang, there is no 'metaphilosophical mandate' demanding that metaphysics takes priority over political theory. This exposes a vulnerable link in the argumentation of Van der Deijl and Smilansky. A political theorist, let's call him McY might simply decide to reject PT, shrugging Van der Deijl's and Smilansky's works off. The argument might be something like this:

1. If PT has not been justified, it is arbitrary.
2. If PT is arbitrary, so are arguments that presuppose it.
3. If these arguments against (my preferred flavor of) egalitarianism are arbitrary, I can offset them by introducing an arbitrary criterion of my own (e.g. conformance to moral intuitions, justification by divine revelation)

Such an argument would, of course, not be very strong in itself, because it merely shifts the burden of proof. That PT is not justified does not necessarily mean it is arbitrary, since a justification might still be found

(contra 1); furthermore, the lack of explicit argumentation for PT does not come close to establishing the contrary (contra 3), regardless of whether that contrary is conceived as the independence or precedence of political theory. This argument only shows that PT is vulnerable; in order for McY's argument to carry any substantive weight, McY would have to argue that his criterion is more attractive than PT, or at least as attractive.

One way to do that might be to capitalize on the fact that PT may have undesirable consequences.

Firstly, acceptance of PT implies that some debates in political theory are impotent until metaphysical issues have been resolved (Fleurbaey, 2001). On the topic of luck egalitarianism, Knight mentions the demarcation of responsibility as an example of such a debate (2006: 814): accepting PT implies that luck egalitarianism can only be applied once consensus has been reached on a metaphysical notion of responsibility. Lacking such an answer, luck egalitarianism would not be able to progress. More generally, PT seems to suggest that we can only engage in political philosophy once we have settled all relevant metaphysical debates.

Secondly, PT has some counterintuitive consequences in the case of free will. Consider the following exchange: a defender of PT responds to McY's argument by holding that it would be inconsistent to use the concept of 'free will' in our political philosophy if our favored metaphysical theories suggest that free will does not exist.<sup>2</sup> McY then brings out the counterintuitiveness of PT with a variation: would it not be weird to do political philosophy without a notion of free will, since our moral intuitions about free will are so 'clear and distinct'? Why would I *a priori* reject a feeling that is so prominently present in every action I undertake? From this perspective, the problem is not that the justification of PT is weak. The problem is, instead, that such a weak justification is used to disqualify moral intuitions and beliefs that play an important part in our life. Even those who support the determinist thesis cannot deny that individuals have the feeling that they are actors that make decisions of their own. Much in the same vein, people who believe in hard determinism can nevertheless have strong moral intuitions about what is fair and what is not. Revising our political theories to comply with PT would classify a large class of appealing moral beliefs as irrelevant, because they cannot be grounded in our metaphysical theory.

Both consequences come together in a corollary of Van der Deijl's argument. If substantive responsibility cannot be saved by hard determinism or compatibilism, and the libertarian position is sufficiently implausible to not be considered, then adherence to PT requires us to reject all positions that presuppose substantive responsibility. This means that individuals cannot, under any circumstance, be seen as making autonomous choices. The consequences of that position are far-reaching. While this is not the place to discuss which positions and domains would be affected, the remainder of this section will try to give some examples.

Isaiah Berlin (2000) suggested that rejecting free will would force us to change the way we talk about human relations:

"Such expressions as 'I should not have done *x*', 'How could you have chosen *x*?' and so on, indeed the entire language of the criticism and assessment of one's own and others' conduct, would undergo a sharp transformation, and the expressions we needed both for descriptive and for practical-corrective, deterrent, hortatory purposes (what others would be open to a consistent determinist?) would necessarily be vastly different from the language which we now use." (Berlin, 2000: 108)

As a part of this change in vocabulary, accepting hard determinism would turn praise and blame to 'purely corrective or educational instruments, or confine them to aesthetic approval or disapproval' (Berlin, 2000: 102). For Berlin, praise and blame do not refer to the kind of praise and blame that can be attributed to someone under Scanlon's attributive responsibility, but only to praise and blame for autonomous actions: one is praised for the autonomous decision to save a drowning child, and one is blamed for the autonomous decision to steal. This kind of praise and blame cannot exist for hard determinists, since every action that is potentially praise- or blameworthy necessarily follows from preceding physical states. The only sense in which we can still speak of praise and blame is, according to Berlin, *aesthetic*. This implies that we can praise someone's bravery only in the same way that we praise someone's beautiful eyes. Both are, after all, the necessary consequence of physical causes; or conversely, neither follows from the autonomous choices of the individual.

What we have lost along the way is a *moral* perspective on praise and blame. As the citation foreshadowed, the loss of this kind of moral praise and blame seems to render statements and questions like ‘I should not have done  $x$ ’ or ‘How could you have chosen  $x$ ?’ empty. If I cannot make autonomous choices, my choice at moment  $t$  was ‘necessary’ in the same way that my regret at moment  $t+1$  is; similarly, it seems to make little sense to blame you for choosing  $x$  if it was not your autonomous choice to begin with.

This point has a practical bearing on our justice system. Courts of law are often confronted with questions of responsibility. Is someone who committed a murder while sleepwalking responsible for his crime, even though he was not conscious (Lamme, 2011)? Is kleptomania an adequate excuse for theft, or should compulsion not count as a mitigating circumstance? If we reject autonomous choices on a metaphysical level and insist, through PT, that our political theory conforms to our rejection of the possibility of autonomous decision making, then our current way of treating these cases needs to be revised.

These examples offer insight into some of the consequences that the rejection of substantive responsibility may have. If we apply PT consistently, much more is lost than particular forms of egalitarianism. Of course, a defender of PT might decide to bite the bullet here. It just seems like quite a big bullet to bite, especially given the fact that PT is not independently justified.

#### 4. John Rawls: the Plausibility of Political Theory without the Posteriority Thesis

This section shows that it is possible to do political philosophy without the Posteriority Thesis and that, consequently, metaphysical theory is not the only legitimate starting point of political philosophy. The lead is taken from Kang (2003). Arguing against Smilansky, Kang held that Rawls has shown how contradicting social values, such as equality and inequality, can be mediated by non-metaphysical criteria. While Rawls defends a limited amount of inequality, he does not do so on metaphysical grounds; that is, because a limited amount of inequality is entailed

by a metaphysical theory about the nature of human beings. Instead, he refers to a category of principles Kang calls ‘rational precepts’ (2003: 117). Examples of such rational precepts are economic analyses and instruments such as Pareto efficiency. Some limited level of inequality can be defended, for example, because it makes everyone better off economically. In deciding what level of inequality to allow in society, Rawls delegates decision-making to procedures on which society has, implicitly, agreed – not to metaphysical considerations.

Kang’s point is that there is more leeway between metaphysics and political theory than Smilansky is willing to admit. However, if we decide to take our cue from Rawls, there might be more lessons to be learned.

In *Justice as Fairness* (2001), Rawls emphasizes that his earlier works were often misinterpreted as built on the foundation of a particular set of philosophical assumptions about the nature of human beings. He argues that his views are to be interpreted as a political, rather than a philosophical theory.

‘Justice as fairness hopes to put aside long-standing religious and philosophical controversies and to avoid relying on any particular comprehensive view. It uses a different idea, that of public justification, and seeks to moderate divisive political conflicts and to specify the conditions of fair social cooperation between citizens.’ (Rawls, 2001: 29)

Rawls’ starting point differs from that of Smilansky and Van der Deijl. Smilansky and Van der Deijl inquire what political theory can be justified on the basis of a particular metaphysical theory; Rawls accepts a pluralist society, in which people have different norms, values and beliefs about the world, as his starting point. Rawls’ outset is to build a bridge between these different moral, religious and philosophical doctrines, rather than to convince anyone that a particular doctrine is wrong and should be exchanged for a different one. The ‘rational precepts’ that Kang refers to are one of the building blocks with which the bridge is built.

It is also interesting to note that Rawls explicitly states that the relevance of free will to political theory is to be drawn from its implicit use in the public sphere. If the metaphysical debate on free will has any importance for political theory, it is relevant because it influences the political views of individuals. Rawls thereby avoids having to deal with difficulties of the kind raised by Van der Deijl and Smilansky.

Richard Rorty's discussion of Rawls' views on the relation between politics and philosophy helps to bring out that Rawls' approach to political theory can coexist with hard determinism, compatibilism and libertarianism alike.

According to Rorty, Rawls' ability to cater to a pluralist society is made possible by the separation of metaphysics and political theory. For Rawls, questions regarding human nature are not relevant when we talk about the organization of our society. Rorty adds a historical twist to Rawls, suggesting that the beliefs and values that Rawls calls 'meanings embedded in political culture' can be seen in a historical and sociological perspective (Rorty, 2008: 181).

This emphasis on the historical contingency of the beliefs and values employed by Rawls should chime well with those who reject free will. Rorty points out, after all, that Rawls does not presuppose that these beliefs and values are the result of autonomous individual agency. He only requires them to be there. In the same vein, this emphasis on the historical contingency of beliefs and values should not prevent a Rawlsian framework from being accepted by compatibilists and libertarians.

The key point is that Rawls' view does not depend on a conception of individuals as either autonomous or not autonomous. This allows Rawls to be genuinely indifferent towards the metaphysical issue of free will: its outcomes do not matter to his method, because his starting point consists of moral dispositions *as experienced*, not *as caused*. This point is important, because it shows that Rawls' theory does not disagree with any specific position in the metaphysical debate – Rawls simply does not need them.

It is perhaps worthwhile to make a short detour to point out that I stray from Kang's own classification of Rawls when I say that he rejects PT. According to Kang, Rawls conforms to PT, but I think this is because

Kang applies his own criteria inconsistently. When discussing Rawls, Kang holds that Rawls *accepts* PT but *deviates* from it when there are rational precepts that justify doing so. Accordingly, Rawls *deviates* from the principle of income equality that follows from his hard determinist intuitions because rational precepts dictate that allowing limited inequality makes everyone better off. At the same time, PT was coined by Kang to represent the view that one's political theory should conform to one's metaphysical theory. In that case stating that Rawls *deviates* from PT when there are rational precepts that justify it seems to be contradictory. The rational precept that Kang uses to illustrate his point is, after all, an ethical statement: inequality is asserted to make 'everyone better off'. I contend Kang presents Rawls as deviating from the thesis that *metaphysics precedes ethics* on *ethical* grounds, which renders the thesis vacuous. PT would then hold that metaphysics precedes ethics unless there are ethical reasons to do otherwise.

That Kang's ascription of PT to Rawls renders the term vacuous can also be seen in a different way, since Kang offers no directions as to what can pass for a 'rational precept'. Kang admits that this question still needs to be answered, but might underestimate the consequence that it has for his categorization of Rawls. If the answer to the question is only a light constraint, then PT is more often ignored than upheld – and *vice versa* if the constraint is more severe. The only interpretation of PT that would be compatible with Rawls is a vacuous one, and consequently I interpret Rawls as rejecting PT.

It is now time to take all this back to the question of this paper. There are two points to be made here. Firstly, on a more general note, the example of Rawls shows that it is possible to justify a political theory on non-metaphysical grounds – in other words, that it is possible to engage in political theory if we do not uphold PT. In the case of Rawls, this is done by taking moral beliefs and dispositions as our starting point.

Secondly, relating specifically to Smilansky and Van der Deijl, the plausibility of a non-metaphysical justification of political theory calls into question whether they were right to dismiss a political theory that could not be justified metaphysically. It is, after all, possible to justify a political theory without any reference to metaphysical theories.

## 5. Conclusion

The aim of this paper has been to evaluate the Posteriority Thesis, as the claim that metaphysical theories take precedence over political theories so that only political theories that conform with a plausible metaphysical theory are admissible. PT is an implicit premise in Smilansky and Van der Deijl's arguments. In this paper I argued that it is questionable whether political theories can be decisively dismissed by such arguments.

The argument proceeded in two steps. Firstly, it was shown that PT itself is currently not supported by an independent argument, and that it has some significant undesirable consequences. Secondly, it was shown that political theories can be legitimately supported by non-metaphysical starting points. The example of Rawls illustrates that consistency with a particular metaphysical theory is not necessarily an attribute of a political theory, weakening the case for rejecting political theories on exclusively metaphysical grounds.

While the arguments of Smilansky and Van der Deijl hold when assuming that one's political theory must conform to one's metaphysical theory, this paper has tried to build credibility for the rejection of that assumption. The main aim of this paper would therefore be achieved when you have been convinced that the Posteriority Thesis is not a trivial assumption. Not only are its consequences counterintuitive and possibly paralyzing for political theory, but unquestioned acceptance of the Posteriority Thesis also risks closing us off from alternative approaches. There is more between heaven and earth than metaphysics.

## Acknowledgements

I am grateful to Joris Vincent, Huub Brouwer, Tom Goodsir, Ingrid Robeyns and the anonymous reviewers of the ESJP for helpful commentary on a draft version of this paper.

*Manuel Buitenhuis holds a Bachelor's degree from University College Utrecht, where he mainly studied philosophy and economics. He currently attends the Erasmus Institute for Philosophy & Economics (EIPE). He also follows a master programme in the History and Philosophy of Science at Utrecht University.*

*'On Metaphysical Cases against Political Theories' is an edited version of a paper that was originally written for the EIPE course 'Ethics & Economics' by prof. dr. Ingrid Robeyns.*

## Notes

1. There is a lot more to be said on this subject, but this definition should suffice for present purposes.
2. Van der Deijl exemplifies this when he holds that 'The point of drawing on arguments from the free will debate for the distributive justice discussion is exactly that it may show us that some of our intuitions may be wrong.' (Van der Deijl 2013: 28).

## References

- Berlin, I. (2000) 'From Hope and Fear Set Free'. In: Berlin, I., *The Proper Study of Mankind: an Anthology of Essays*, 91-118. New York: Farrar, Straus and Giroux.
- Dworkin, R. (2002) *Sovereign Virtue: the Theory and Practice of Equality*. Cambridge: Harvard University Press.
- Fleurbaey, M. (2001) 'Egalitarian Opportunities'. In: *Law and Philosophy* 20(5), 499-530.
- Kang, S.-H. (2003) 'Free Will and Distributive Justice: A Reply to Smilansky'. In: *Philosophia* 31(1-2), 107-126.
- Knight, C. (2006) 'The Metaphysical Case for Luck Egalitarianism'. In: *Social Theory and Practice* 32(2), 173-189.
- Lamme, V. (2011) *De Vrije Wil Bestaat Niet*. Amsterdam: Uitgeverij Bert Bakker.
- Rawls, J. (2001) *Justice as Fairness*. Cambridge: The Belknap Press of Harvard University Press.

- Rorty, R. (2008) 'The Priority of Democracy over Philosophy'. In: Rorty, R., *Objectivity, Relativism, and Truth (Philosophical Papers: Volume 1)*, 175-196. New York: Cambridge University Press.
- Scanlon, T. (1999) *What We Owe to Each Other*. Cambridge: The Belknap Press of Harvard University Press.
- Smilansky, S. (1997) 'Egalitarian Justice and the Importance of the Free Will Problem'. In: *Philosophia* 25(1-4), 153-161.
- Van der Deijl, W. (2013) 'The Metaphysical Case against Luck Egalitarianism'. In: *Erasmus Student Journal of Philosophy* 4, 22-30.

