APORIA

JOURNAL OF THE PHILOSOPHY SOCIETY OF ST ANDREWS

Vol. XVI

ABOUT

Aporia is an established undergraduate Philosophy journal that has been published by the University of St Andrews Philosophy Society since 2007. It is edited by Kirsty Hardwick (editor in chief) and Alice Fanner. *Aporia* serves as a reputable publishing platform for undergraduates and recent graduates worldwide.

Aporia is funded by the University of St Andrews Philosophy Society, which receives funds from the University of St Andrews Department of Philosophy, the University of St Andrews Union, and independent benefactors.

ACKNOWLEDGEMENTS

The editors would like to thank both individuals and institutions that continue to support the journal, including philosophy departments and societies worldwide for distributing the journal's Call for Papers; individuals who submitted papers; the institutions who fund the society; and the University of St Andrews. Special thanks also go to Oscar Westerblad and Quentin Pharr.

Contents

Presentism, Temporal Distributional Properties, and Fundamentality*

Matthew Green[1]

Abstract

According to presentism, everything that exists is present. According to the truthmaker principle, for every true proposition there is a truthmaker – an entity that suffices for the truth of that proposition. According to realism about the past, there are true propositions about the past. Together these claims necessitate presently existing truthmakers for truths about the past (presentist truthmakers).

Cameron (2010) argues that temporal distributional properties (TDPs) can play the role of presentist truthmakers. Corkum (2014) argues that they cannot. I argue against Corkum's objections. In §2, I introduce, and outline the motivation for, TDPs. In §3, I show that unless TDPs are stipulated to be fundamental, as Cameron does, they can be reduced to temporal non–distributional properties, which are unable to play the role of presentist truthmakers. In §4, I argue against Corkum's two objections to Cameron's stipulation. Corkum's first objection is that Cameron has no grounds on which to stipulate that TDPs are fundamental, and that the reducibility of TDPs to temporal non–distributional properties (as discussed in §3) shows that they are not. I argue that the burden of proof is not on Cameron to argue that TDPs are fundamental, but on Corkum to argue that they are not, and that to argue from the reducibility of TDPs to their non–fundamentality is to beg the question against Cameron: the reduction is only possible once their non–fundamentality is assumed. Corkum's second objection is that if Cameron is allowed to stipulate that TDPs are fundamental in order to escape objections, then a superior alternative account is allowed to make the same move, rendering Cameron's account redundant. I argue that the cases are asymmetric: the alternative account faces a legitimate objection whilst Cameron's account does not.

1 Introduction

According to presentism, everything that exists is present. According to the truthmaker principle, for every true proposition there is a truthmaker – an entity that suffices for the truth of that proposition. According to realism about the past, there are true propositions about the past. Together these claims necessitate presently existing truthmakers for truths about the past (presentist truthmakers). Cameron (2010) argues that temporal distributional properties can play the role of presentist truthmakers. Corkum (2014) argues that they cannot.

* Thanks to Ross Cameron, Simon Hewitt, Robin LePoidevin, Helen Steward, and Jason Turner for useful comments on numerous drafts of this essay.

[1] Matthew Green has a BA and MA in Philosophy from the University of Leeds, and is currently taking an MLitt in Philosophy at the University of St Andrews. So far, he has worked mainly in philosophy of language, philosophy of logic, and metaphysics.

In this essay, I argue against Corkum's objections. In §2, I introduce, and outline the motivation for, temporal distributional properties. In §3, I show that unless temporal distributional properties are stipulated to be fundamental, as Cameron does, they can be reduced to temporal non–distributional properties, which are unable to play the role of presentist truthmakers. In §4, I argue against Corkum's two objections to Cameron's stipulation. Corkum's first objection is that Cameron has no grounds on which to stipulate that temporal distributional properties are fundamental, and that the reducibility of temporal distributional properties to temporal non–distributional properties (as discussed in §3) shows that they are not. I argue that the burden of proof is not on Cameron to argue that temporal distributional properties are fundamental, but on Corkum to argue that they are not, and that to argue from the reducibility of temporal distributional properties to their non–fundamentality is to beg the question against Cameron: the reduction is only possible once their non–fundamentality is assumed. Corkum's second objection is that if Cameron is allowed to stipulate that temporal distributional properties are fundamental in order to escape objections, then a superior alternative account is allowed to make the same move, rendering Cameron's account redundant. I argue that the cases are asymmetric: the alternative account faces a legitimate objection whilst Cameron's account does not.

2 Intrinsic Determination and Temporal Distributional Properties

One attempt at providing the presentist with a truthmaker for past truths is Bigelow's (1996) account of Lucretian properties – presently existing tensed properties instantiated by the world. Under Bigelow's account, for example, 'dinosaurs once existed' is made true by a presently existing tensed Lucretian property instantiated by the world – something like <being such so as to have contained dinosaurs>.

Sider (2001) charges Lucretian properties with ontological cheating, by which he means two things: they are hypothetical (they 'point–beyond' their bearers) and they are not categorical (they do not 'point–at' their bearers). To quote Sider (2001, 41), "[w]hether the world has the property *previously containing dinosaurs* is not a matter of what the world itself is like, but points beyond itself, to the past." Cameron (2010) takes these charges to amount to the claim that Lucretian properties tell us nothing about the intrinsic nature of their bearers in the present, and proposes to admit only difference–making properties – properties that make a difference to a bearer's intrinsic nature – into our ontology, sifting out unwanted non–difference–making properties with the criterion of intrinsic determination:

> for all objects x and properties F and times t, if x instantiates F at t, then x has the intrinsic nature at t that it has partly in virtue of instantiating F at t.
>
> (Cameron, 2010, 5)

Bigelow's properties do not satisfy intrinsic determination: the world does not have its present intrinsic nature (even partly) in virtue of instantiating the property <*being such as to have contained dinosaurs*> (another world could have a different history to our own and thus fail to instantiate this property, yet nonetheless be identical to our own world in terms of its present intrinsic properties). So, Lucretian properties cannot serve as presentist truthmakers.

Cameron's claim is that temporal distributional properties do the job of providing for presentist truthmakers. A distributional property is the way that something is across a region: a spatial distributional property is the way that something is across space; a temporal distributional property is the way that something is across time. As Josh Parsons (2004, 1) puts it, "a distributional property is like a way of painting, or filling in, a spatially [or temporally] extended object with some property". For example, 'dinosaurs once existed' is made true by a presently existing temporal distributional property instantiated by the world, which details how the world has been at every instant in time – something like *<being like x, then like y>*[2], where x maximally describes the world at some instant of time in which it contained dinosaurs. It is important to note that it is not just one part of the property, the *<being like x>* part, that serves as the truthmaker for 'dinosaurs once existed'. Temporal distributional properties (and properties more generally) do not have parts. It is the property as a whole which serves as the truthmaker, which is not divided but spread across its bearer.[3]

Temporal distributional properties satisfy intrinsic determination: the world does have its present intrinsic nature in virtue of instantiating *<being like x, then like y>*, where y maximally describes the world right now, so they are difference making. Again, this is the case not because *<being like y>* is a part of *<being like x, then like y>*, but because of the latter property as a whole. All else being well, then, temporal distributional properties seem fit to serve as presentist truthmakers.

3 The Reducibility of Temporal Distributional Properties

The reductionist about distributional properties thinks that distributional properties like*<being polka–dotted red on white>* can be reduced to non–distributional properties like *<being an x such that there are some ys, and the ys are part of x, and the ys are of the right sorts of colour, and the ys are spatially related in the right sorts of ways>*.[4] Parsons (2004) objects to the possibility of such a reduction by pointing out that non–idealised (i.e., real world) instances of polka–dots will have a non-uniform colour distribution that cannot be reduced to a non–distributional colour property. No non–idealised polka dot will be uniformly red, and that heterogeneity will make it impossible to characterise it as a y of the right sort of colour. Moreover, the reductionist has to contend with gradients of colour, from red to blue, say, and here again there will be no ys that are the right sort of colour, for any extended y will be a gradient of some colour to another between red and blue (perhaps some very closely related purples).

To avoid this problem, the reductionist must substitute polka–dots for extensionless points which can bear neither non–uniform colour distributions nor, more generally, distributional properties. Parsons puts it like this:

> [W]hen, and only when, red and white point–sized objects [i.e., the smallest spatial parts of the property bearer] are spatially arranged in the right way, you get something polka–dotted red on white. [Then]

[2] More accurately, *<being like . . . , then like x, then . . . , then like y, then . . . >* for every way the world has been, is, or will be; but that's stylistically unwieldy, so I'll use the formulation *<being like x, then like y>* to mean this.

[3] Thanks to Helen Steward for insisting on clarity here.

[4] Example properties due to Parsons (2004, 5)

> [<*being polka–dotted red on white*>] is equivalent to a non–distributional property specifying the arrangement of such points, and the same sort of story will work for any other distributional property.
>
> (Parsons, 2004, 5)

If the reduction is to be warranted in general then this must be possible in every case, and so every extended object must consist of extensionless points (for if it didn't there would be room for distributional properties). This places a restriction on the reductionist that she cannot meet: given the possibility of gunk (whereby there are no points, that is, space is infinitely divisible) or spatially extended simples (whereby a partless object occupies two points or more, that is, exists (simpliciter) at more than one point), there is room for an extended object with no extensionless parts and so room for a spatial distributional property that will not reduce (Parsons, 2004, 6-8). That gives us reason to rule out the equivalence of the two kinds of property, and so reason to rule out the reduction.

The same argument against reduction does not hold for temporal distributional properties; at least not for presentists. Consider the temporal distributional property <*being like x, then like y*>. Just as the property <*being polka–dotted red on white*> is distributed across its spatially extended bearer, so too the property <*being like x, then like y*> is distributed across its temporally extended bearer, the world. To paraphrase Parsons:

> [W]hen, and only when, x–like and y–like instance–sized objects [i.e., the smallest temporal parts of the property bearer] are temporally arranged in the right way, you get something x–like then y–like. Then <*being like x, then like y*> is equivalent to a non–distributional property specifying the arrangement of such instants, and the same sort of story will work for any other temporal distributional property.

If the reduction is to be warranted in general then this must be possible in every case, and so every temporally extended object must consist of extensionless instants. This places a restriction on the reductionist that she can meet, but only if she's a presentist. Given presentism there is no possibility of temporal gunk. If there were, then any time which is present would be divided into smaller parts, all of which would then be present. But the present exists only at an instant. So, there is no temporal gunk. There is also no possibility of temporally extended simples. If there were, then they would exist (simpliciter) at more than one time. But, according to presentism, everything that exists exists in the present, and as we have just seen, the present cannot occupy more than one time. So, there are no extended simples. Thus, there is no room for an extended object with no extensionless temporal parts, and so no room for a temporal distributional property that will not reduce. That demonstrates the equivalence of the two kinds of property, and that warrants the reduction.

As such, there is (yet) no principled metaphysical reason why temporal distributional properties cannot be reduced to conjunctive non–distributional properties. Let's call this the reduction observation. This is of potential concern to Cameron, since conjunctive non–distributional properties reduced from temporal distributional properties have parts that are wholly past–facing. For example, the temporal distributional property <*being like x, then like y*> is reduced to a conjunctive non–distributional property, the <*being like x*>) conjunct of which is wholly past–facing, and which thus fails to satisfy intrinsic determination, ruling it out as a legitimate property, and ruling it out as a presentist truthmaker.

Cameron's reply is to block the reduction by stipulating that temporal distributional properties are fundamental:

> It's true that for any temporal distributional property, there's a very complex big conjunctive property made up of the objectionable past-directed properties and unobjectionable present-directed properties [the] instantiation of which would have exactly the same effect as the distributional property. In that sense, they're equivalent. But the temporal distributional property is fundamental, whereas the conjunctive property is not. [...] Distributional properties cannot be broken up into simpler components: there is just one property here, and it is fundamental – and it is exactly the same property that is grounding truths about how the bearer now is that is grounding truths about how the bearer was.
>
> (Cameron, 2010, 11)

Since temporal distributional properties are stipulated to be fundamental, they cannot be reduced to non–distributional properties, leaving it open to the presentist to employ them as truthmakers.

4 The Fundamentality of Temporal Distributional Properties

Corkum's claim is that Cameron's stipulation of the fundamentality of temporal distributional properties is illegitimate. His first objection has two parts: first, that Cameron has no grounds on which to stipulate that temporal distributional properties are fundamental, and second, that the reduction observation is evidence that they are not.

> On what grounds can the distributionalist make this stipulation? It can be difficult to adjudicate on such matters. One philosopher's bold conjecture is another's undefended assumption. However, in this case the onus is clearly on the distributionalist to defend the stipulation. Some of our examples of temporally distributed properties appear to be conjunctive properties and so prima facie not irreducible. For example, the property of having been a child and being now an adult appears to be explicable in terms of the properties of being a child and being an adult, along with a temporal arrangement of these properties
>
> (Corkum, 2014, 5)

With regard to the first part, Cameron's reply is that the proponent of a theory is permitted to stipulate anything they want.[5] Of course, once something has been stipulated it can be objected to (the objection can hardly be made prior to that which is being objected to), but then the burden of proof lies with the opponent. Sider (2006, 81-82) makes the point that the proponent of a thesis is under no obligation to convert his opponent to his way of thinking, and is required only to resist her attacks. That applies here: it is perfectly legitimate for Cameron to stipulate that temporal

[5] Thanks to Ross Cameron, Helen Steward, and Jason Turner for convincing me of this point.

distributional properties are fundamental, and if his opponent doesn't think that they are fundamental, then the burden of proof is on her to demonstrate why.

With regard to the second part, I take Corkum to be saying that even if it were legitimate for Cameron to initially stipulate the irreducibility of temporal distributional properties without any evidence for their fundamentality, this changes when charges are bought against it. The reduction observation is evidence against the fundamentality of Cameron's property, and the burden of proof is now on Cameron to provide evidence for their irreducibility, which he cannot.

But that's not quite right. The initial stipulation of fundamentality blocks that reduction. As Sider pointed out, Cameron is under no obligation to convert Corkum, and Cameron is under no obligation to provide evidence for the fundamentality of his properties. His intellectual obligation in the face of Corkum's objections is not to provide arguments for the thesis that temporal distributional properties are fundamental, but rather to provide arguments against those objections – against the reductionist thesis that they are explicable in terms of non–distributional properties appropriately arranged. And this he has done by stipulating their fundamentality: temporal distributional properties are not explicable in terms of non–distributional properties because the latter have parts that are wholly past–facing and as such fail to satisfy intrinsic determination, and the former do not. You can't explain one thing in terms of another thing when they're demonstrably inequivalent.

So far, then, I've argued that it is legitimate for Cameron to stipulate the fundamentality of temporal distributional properties, and that, once stipulated as such, the reduction observation does not constitute an objection against their fundamentality.

Corkum's second objection is that if it is legitimate for Cameron to stipulate that temporal distributional properties are fundamental, then it is legitimate for proponents of Lucretian properties to stipulate that Lucretian properties satisfy intrinsic determination. If this is true, then either (i) if the consequent is false, then the antecedent is false by modus tollens, or (ii) if the consequent is true, then there is no need for Cameron's properties, so his stipulation of their fundamentality is self–defeating.[6]

I take Corkum's reasoning to be as follows. The reduction observation gives us reason to believe that temporal distributional properties reduce to non–distributional properties, and if that is the case, then they can't do the work that presentists require of them. In response, Cameron just stipulates that they don't reduce, and the problem goes away. Analogously, there is reason to believe (see §1) that Lucretian properties do not satisfy intrinsic determination, and if that is the case, then they can't do the work that presentists require of them. So in response, why can't Bigelow just stipulate that they do satisfy intrinsic determination, and let the problem go away? If Cameron can do it, why can't the proponent of Lucretian properties?

[6] I'll leave aside discussion of the claim that the rehabilitation of Bigelow's properties would make Cameron's redundant. All Corkum has to say on the matter is this: "But if it is permissible to stipulate that Lucretian properties [satisfy intrinsic determination], then the Lucretian can meet the challenge [of providing for a presentist truthmaker], and there is no reason to embrace [temporal distributional properties]. For these reasons, the stipulation that temporal distributional properties are [fundamental] is either unpersuasive or self-defeating."

To this, Corkum anticipates the following response:

> The distributionalist may respond that there is a salient difference between the two cases. In the one case, Lucretian properties are stipulated as making an intrinsic difference in the world; in the other case, temporal distributional properties are stipulated to be irreducible.
>
> (Corkum, 2014, 28)

Corkum is right that the distributionalist will respond by pointing out a salient difference in the cases, but he misidentifies what that salient difference is. What is being stipulated is irrelevant. What's relevant is that whilst there is an argument forwarded for the thesis that Lucretian properties do not satisfy intrinsic determination, there is no argument forwarded for the thesis that temporal distributional properties are not fundamental. The salient difference is that there are arguments against the satisfaction of intrinsic determination by Lucretian properties where there are none against the fundamentality of temporal distributional properties. One might think that the reduction observation counts as just such an objection, but that objection would beg the question: the reduction only goes through if temporal distributional properties are non–fundamental, and to assume that they are in order to prove that they are is question begging.

Why is this the salient difference? Well, assume that in the formulation of their respective theses, the proponent of Lucretian properties stipulates that they satisfy intrinsic determination, and the proponent of temporal distributional properties stipulates that they are fundamental. Now assume that an objection is given to each property. Clearly, in both cases, reiterating the initial stipulation is not enough, as it is still subject to the objection raised. What is required is a defence of the stipulation against the attack. In actuality, just such an objection has been raised for the proponent of Lucretian properties, and it is obviously not enough just to reiterate the claim that is under attack. No such objection has been raised for Cameron, so his initial stipulation remains unsullied. Put another way, the salient difference is this: Cameron's stipulation is a legitimate initial formulation; Corkam's proposed stipulation is an illegitimate response to objections.

5 Conclusion

Cameron's claim is that temporal distributional properties are truthmakers, which, if true, resolves the trilemma between presentism, truthmaker theory, and past truths. To do so he needs to stipulate that they are fundamental, which Corkum objects to. I have argued that Corkum's first objection fails because (i) it is legitimate for Cameron to stipulate their fundamentality, and (ii) given that stipulation, the reduction observation fails as an objection to their fundamentality. I have also argued that Corkum's second objection fails because, whilst there is an objection made against the satisfation of intrinsic determination by Lucretian properties, there is no objection made against the fundamentality of temporal distributional properties. Of course, whether or not we are justified in believing that temporal distributional properties are truthmakers under presentism depends, in part, on whether or not we are justified in believing that they are fundamental. My claim here is not that they are, but that Corkum's argument that they are not is unsuccessful. It remains to be seen whether there are such arguments.

Bibliography

Bigelow, John. "Presentism and Properties." *Philosophical Perspectives 10* (1996): 35-52.

Cameron, Ross. "Truthmaking for Presentists." *Oxford Studies in Metaphysics 6* (2010): 55-100. (page numbers from: http://myweb.facstaff.wwu.edu/wasserr/bspc/underground/Cameron.pdf).

Corkum, Phil. "Presentism, Truthmakers, and Distributional Properties." *Synthese 191* (2014): 3427-46. (page numbers from: http://www.ualberta.ca/ pcorkum/presentism.pdf).

Parsons, Josh. 'Distributional Properties', in *Lewisian Themes: The Philosophy of David K. Lewis,* edited by Frank Jackson, and Graham Priest*,* Oxford: Oxford University Press, 2004.

Sider, Theodore. *Four–Dimensionalism: An Ontology of Persistence and Time,* Oxford: Clarendon Press, 2001.

Sider, Theodore."Quantifiers and Temporal Ontology." *Mind 115* (2006): 75-97.

A Remark on Attitudes De Dicto and De Se

Shuguo Tang[1]

Abstract

This is a two-fold response to the theory of attitudes presented by David Lewis in his paper *Attitudes De Dicto and De Se*. Lewis' theory is on the specification of attitudes. On the one hand, it is obvious in his text that Lewis attempts to specify the objects of attitudes uniformly. I will argue that the attempt fails. The reason is that from Lewis' formulation of an attitude, we cannot read the object of it. On the other hand, Lewis' account could have been a theory on the specification of attitudes without any commitment on what the objects of attitudes are. I will argue that this does not hold either. In order to do that, I will construct pairs of cases in which the attitude in one case has the same formulization as in the other case according to Lewis' theory. I will argue that in one such pair of cases, the *focus* of the attitude in one case can be different from that in the other, while the Lewisian theory of attitudes does not help to specify an attitude's *focus*.

1 Introduction

In his paper *Attitudes De Dicto and De Se*, David Lewis gives an account which claims that what an attitude is about can be specified uniformly in terms of properties. In this paper, I attempt to challenge Lewis' account.

We have good reasons to declare Lewis' success if his account establishes that properties should be treated as objects of attitudes. However, I will argue that his attempted establishment does not work. At the first stage of my challenge, I will examine Lewis' explicit claim that it is better to hold that objects of attitudes are properties than that they are propositions (Lewis, p.514). I will argue that Lewis' analysis does not exclude the possibility that objects of attitudes are something other than properties.

2 Lewis on why properties, not propositions, are the objects of attitudes

In the following discussion, I will adhere to Lewis' notion that a proposition can be identified with the set of all possible worlds where this proposition is true. For example, we can imagine possible worlds in which grass is red and possible worlds in which grass is green. The proposition 'grass is green' is therefore the set of all possible worlds in which grass is green.

It seems that objects of attitudes are not uniform. Take desire as the paradigm attitude. I want a Shiba Inu and a good grade for my mathematics class. It does not make sense to place a Shiba Inu and getting a decent grade into one single category. So, if objects of desires are uniform, then my object of desire had better not be a Shiba Inu or getting a decent grade. One group of popular candidates for objects of

[1] Shuguo Tang is a third-year undergraduate student of philosophy at Magdalen College, Oxford. He is working on free will, propositional attitudes, and Kant's conception of natural laws.

attitudes would be propositions. When I want a Shiba Inu, what I want is that 'I own a Shiba Inu' be true. When I want to get a decent grade, what I want is that 'I get a decent grade' be true. Apply the same line of paraphrasing mutatis mutandis to other attitudes like beliefs and knowledge, now it seems that my objects of desires are pretty much in order. Yet, despite the seeming versatility of propositions, Lewis claims that they do not capture all conceivable objects of desires and that properties are better candidates. Lewis takes two steps toward this claim. First, he argues that properties can account for at least as many cases of attitudes as propositions; and second, there are some cases that cannot be handled by propositions but can be handled by properties (Lewis, p.514).

Lewis observes that each proposition corresponds to a property of inhabiting a world which is a member of this proposition. Therefore, whenever we say one wants that P be true (where 'P' represents a proposition), it does as well to say that she wants to possess the property of inhabiting a world where P is true. Certainly, when I want that the earth be full of Shiba Inus, I want to inhabit a world in which the earth is full of Shiba Inus.

If objects of desires are propositions, then the more consistent propositions one includes in one's desires, the smaller the set of possible worlds satisfying one's desire is, until the conjunction of propositions desired determines a single world. At that time, there is nothing left to be desired. Further, there are no desires left to be satisfied once the character of the world is completely specified. Nevertheless, Lewis argues that there is something other than propositions that can be desired. Lewis asks us to imagine a world in which there are two Gods, one of which lives on the top of the highest mountain and the other which lives on the top of the coldest mountain, neither of which knows on which mountain they are situated. We assume that the desires of the two Gods specify the same class of worlds. However, even though this set includes the desires of the truth of 'such-and-such a God lives on the highest mountain' and 'such-and-such a God lives on the coldest mountain', it does not determine which God wants to live on which mountain. Suppose that one of them suddenly wants to live on the highest mountain, he would not contradict his previous desires, for he wants the same world still to be actual but only wants a specific location in the world. That is to say, this further desire cannot be accounted for by propositions which make the world distinct, but by the property of being located on the highest mountain which is within the world in question. (Lewis, p.529) Likewise, when I want to own a Shiba Inu, I do not merely want to inhabit a world in which someone satisfying the description of Shuguo Tang owns a Shiba Inu, but I also want that I am one of them. Hence, there are objects of desires beyond the scope of propositions.

In this section, I have presented Lewis' motivations for arguing that cases of desires can be specified in terms of properties uniformly. That is to say, a subject S desires X, if and only if the Lewisian formulation of the attitude 'S intends to be assigned such-and-such property Y' holds. For example, when I desire that a person called Shuguo Tang owns a Shiba Inu, I desire to be assigned the property of inhabiting a world in which Shuguo Tang owns a Shiba Inu. But when I desire that I myself own a Shiba Inu, I desire to be assigned the property of owning a Shiba Inu which distinguishes me as having this privilege. However, one might question whether properties as Lewis 'conceives them', suffice as objects of attitudes in the general case. In §3, I take up this objection and argue that Lewis' analysis does not exclude the possibility that objects of attitudes are something other than properties.

3 Can properties account for the focus of an attitude?

Consider this statement: a camera takes a photo A if and only if it catches such and such combination of light L. While it makes sense to say, appealing to the combination of light, that we have a uniform explanation of what goes on when a camera takes a photo, it does not follow that the objects of the photo are of one single kind. Instead, a photo contains the images of many different objects: a pig, a cat, a tower, the sky, the sun, etc. Similarly, the physical process of eating an apple can also be described fully in terms of how atoms of the apple interact with my mouth, but it is the apple that I am eating, not the atoms. In the same way, even though we fully and uniformly describe, as far as Lewis's conception of properties is concerned, what is going on when one is desiring (using a Lewisian formulation of the attitude), the objects of desires may still be abundant. Further, even if we allow that specifications of attitudinal objects supervene on Lewisian formulations of attitudes, we would still need to work out how the former corresponds to the latter[2].

Lewis' account could have been a theory attempting to specify what our attitudes are about without postulating properties as attitudinal objects. In the following, I will argue that, if Lewis insists that a property is the set of objects that instantiate it, such an attempt cannot succeed either (Lewis, p.515).

An attitude has a *focus*. What I call a *focus* is the feature of the attitudinal object on which my attention is directed. A successful theory of what attitudes are about should be able to specify the *focus* of an attitude. Given Lewis' notion of properties, if properties P and Q pick out the same particular object, then their Lewisian descriptions cannot be different, i.e. 'S intends to be assigned P' is the same as 'S intends to be assigned Q'. The problem for Lewis' account is that, in many cases, a pair of properties P and Q can specify the same object while the *focus* specified by the P-formulation is different from that specified by the Q-formulation.

A Lewisian formulation of an attitude, granted that it uniquely corresponds to the attitudinal object, does not necessarily tell you what the attitudinal *focus* is, just as a complete description of a photo (by specifying the relevant combination of light perhaps) cannot tell you the *focus* of creation and appreciation intended. For example, in the same work of photography, the shape of the Statue of Liberty could be the *focus* or the pose of the Statue of Liberty could be the *focus*. Importantly, a complete description of the physical image does not determine what the *focus* of the work is.

For a pair of properties that pick out the same object, I can think of three cases in which it seems that one specifies a different *focus* from that specified by the other. In the first case, one has an attitude towards inhabiting the world in which a necessary proposition is true. For example, John may desire to be assigned the property of living in a world in which 2+2=4. Mary may desire to be assigned the property of living in a world in which 3+3=6. In the second case, the instantiation of one of the properties necessitates instantiation of the other property. For instance, in a paradigmatic example of the second case, John wishes that the triangle in front of him is equilateral while Mary wishes that it is equiangular. Finally, in the third

[2] If A supervenes on B, then B determines A and not vice versa. For example, the room temperature supervenes on the complete status of air molecules in the room which includes each molecule's velocity: given the complete status of air molecules, we can calculate the average kinetic energy for this group of molecules which is equal to the room temperature; but given the room temperature, or the average kinetic energy, we still cannot know what is the complete status of air molecules in the room; e.g. to change the complete status but not the temperature, we can subtract 1J from a molecule's kinetic energy while add to another's kinetic energy 1J.

case, the same figure can be conceptualised as different artworks. For example, a photo of the Statue of Liberty can be appreciated as a good work for *The Shape of the Statue of Liberty*, or it can be highly regarded as *The Pose of the Statue of Liberty*. Under the former title, the *focus* of the audiences' appreciation is drawn to the geometrical patterns of the photo, while the second title guides the audiences to attend to an animate element of the Statue of Liberty. It could be the case that given the same pose, the photo shows the same shape and vice versa. Lewis himself realizes (and promptly dismisses) the problems in the first and the second cases (Lewis, p.515) but since dismissing these cases as marginal is not a satisfying approach, it is worth considering all three cases in more detail.

One might attempt to fix the difficulties in the first and the second cases by appealing to ignorance. In this line of argument, for the first case, if John were ignorant of the co-extensionality of the properties 'inhabiting the world in which 2+2=4' and 'inhabiting the world in which 3+3=6', then John would probably think that Mary and he desired to be located in different worlds; for the second case, again, suppose John and Mary were ignorant about the co-extensionality of the properties 'being an equilateral triangle' and 'being an equiangular triangle'. One could even imagine a quarrel between them on whether an equilateral triangle is better than an equiangular triangle. In other words, the *focus* of John's attitude can be the same as that of Mary's, even though they do not think so. The argument from ignorance not only owes us an explanation to why even a mathematician can have an impression of *focal* difference struck by the first and the second case, but also threatens to isolate knowledge as a special species of attitude. *But* for now, what is interesting about this line of argument, is that, whatever the verdict about the role of ignorance in certain cases, the third case introduced in the past paragraph avoids this concern altogether.

Suppose that John and Mary took a photo of the back of the Statue of Liberty. John and Mary dispute upon whether to name the photo *The Shape of the Statue of Liberty* or *The Pose of the Statue of Liberty*. John argues that looking at the back of the Statue of Liberty, the delicate corrugated shape of the robe stands out in contrast with the relatively plain pose shown, as from this angle we cannot even see her with a book at her left hand. Mary argues that the patterns of the back of the Statue of Liberty are much less abundant than that of the front and such a simple arrangement enables us to appreciate the power of the pose. So, we have the shape-formulation of John's attitude that John appreciates being assigned the property of entertaining the photo of the statue of liberty showing such-and-such shape and a similar pose-formulation of Mary's attitude.

Both shape- and pose-formulations seem to have a potential to specify a different *focus* from the other, but the focal difference cannot be explained by anyone's ignorance of the properties' co-extension. Both John and Mary know perfectly well that the shape-formulation and the pose-formulation necessitate each other. But when John and Mary are admiring the photo, given the arguments by them, we can see that the *focus* of John's attitude is different from that of Mary's.

Different *characterizations* of the Statue of Liberty triggered by 'shape' and 'pose' may help explain and justify why the former specifies a different *focus* from that specified by the latter in the third case. *Characterizations* are potentially very personalised frames in which we attend to objects. For a same piece of cloud resembling an animal, it can be the case that a person attends to it as if it is a horse while another person attends to it as if it is a mule. That is to say, the former person *characterises* the cloud as horse-like while the latter *characterises* it as mule-like. In the case of the photograph, 'Shape' normally helps to trigger a *characterization* of the Statue of

Liberty regarding its geometrical patterns that guides our attention to that geometrical feature that is the *focus* for John's attitude, while 'pose' triggers a *characterization* regarding that animate element that guides our attention to its anthropomorphic feature which is the *focus* for Mary's attitude.

There is no reason to think that different *characterizations* of an object cannot be triggered by other pairs of co-extensional properties. If so, then the *focal* difference in the first and the second case may also be explained and justified. For the first case, we can see the difference between necessary truths 2+2=4 and 3+3=6, since the calculation of the former requires four fingers and that of the latter requires two more. So, the *characterization* of a world triggered by the former fact can be different from that triggered by the latter. One might desire to inhabit a world in which 2+2=4 while 3+3=5 because this person does not want to calculate with more than one hand. However insane, this is a conceivable attitude.

In the second case, for the same triangle, since 'equally long' has a different meaning from 'equally acute', 'equilateral triangle' and 'equiangular triangle' can trigger different *characterizations*. Imagine that John and Mary have a very sharp triangle ABC designed to punch holes on a plain of wood. Here is how it works. The triangle shall be always perpendicular to the wooden plain. Starting with the *segment* AB on the wood, by evenly rotating it in order for AC to land on the plain, they punch a hole during the process with the *angle* BAC; then they turn AC evenly in order for BC to land on the plain and punch a hole during the process with the *angle* ACB, and so on… Now they have a series of holes on the wooden plain. Hoping that the distance between any two adjacent holes is equal, John expresses his wish for ABC to be an equilateral triangle. Hoping that each hole is equally deep, Mary express her wish for ABC to be an equiangular triangle. The *focus* of John's wishing is the situation among angles' acuteness in ABC, while the *focus* of Mary's wishing is the situation concerning the segments' lengths in ABC. However, such *focal* difference which can be triggered by *characterisations* of ABC as 'being equilateral triangle' and 'being equiangular triangle' cannot be captured by Lewis, since keeping the commitment that a property is defined by its extension, Lewis has to say that these two *characterisations* actually represent the same property.

Conclusion

To specify what an attitude is about, we need to know both the object of the attitude and the *focus* of the attitude. While Lewis' formulation of an attitude corresponds to the object, I have argued that it does not necessarily establish that properties should be treated as attitudinal objects. It may be the case that we can in principle pick out specifications of the attitudinal object from Lewisian formulations since the former supervenes on the latter, but it may be a challenge for Lewis to provide us with standardized procedures to do so. Furthermore, since Lewis' notion of property does not distinguish between a pair of co-extensional properties, his account is not sensitive to the *focal* difference between the properties. I conclude that Lewis' account of what attitudes are about is incomplete until it can specify the attitudinal *focus*.

Bibliography

Lewis, David. "Attitudes De Dicto and De Se." *The Philosophical Review 88* (1979): 513-543.

Is intuition best treated as a sui generis mental state, or as a belief?

Joshua Kelsall[1]

Abstract

It is common in philosophy for philosophers to consult their intuitions regarding philosophical issues, and then use those intuitions as evidence for their arguments. For instance, an incompatibilist about moral responsibility might argue that her position is correct because it is intuitive that, given a deterministic world, people cannot be morally responsible. One might ask whether or not the philosopher is justified in using intuitions in her argument, but it seems that in order to answer this, we require an understanding of what intuitions are. One characterisation of intuitions is that they are either ordinary beliefs, beliefs with particular aetiologies, or inclinations to believe. An alternative view is that intuitions are unique, sui generis mental states, unique in the sense that they cannot be fully understood in terms of other mental states. In this essay, I defend the belief view of intuitions. I argue that intuitions are inclinations to believe, which are derived from unconscious reasoning processes which take our background beliefs and theories as input. I defend this view against Elijah Chudnoff's claim that intuitions are sui generis mental states akin to perceptual experiences. I suggest that the analogy Chudnoff draws between perceptions and intuitions does not hold, because perceptions are belief independent, while intuitions are not. I then respond to Chudnoff's counter-example to the belief view, which is that it is possible to have an intuition but no corresponding belief or inclination to believe. I formulate an error theory which demonstrates that it is possible to confuse the rejection of an intuition as part of our consciously derived beliefs, with the elimination of an intuition as part of our unconsciously derived beliefs. Such confusion, I suggest, leads to Chudnoff's putative counter-example that it is possible to have an intuition and yet no inclination to believe in it.

1 Introduction

Intuitions can be roughly described as a kind of experience in which it seems to an agent, that something or other, is or is not, the case. We can ask of these experiences, what kinds of experiences are they? On this, there are two prominent analyses. On the one hand, there is the belief view, according to which intuitions are either ordinary beliefs, beliefs with particular aetiologies, or inclinations to believe. On the other hand is the view that intuitions are unique, sui generis mental states, unique in the sense that they cannot be fully understood in terms of other mental states. In this essay, I defend the belief view of intuition. My aim will be to argue that intuitions are best understood as inclinations to believe, which are caused by unconscious reasoning processes which take our background beliefs and theories as input. I defend this view against Elijah Chudnoff's (2011) opposing claim that intuitions are sui generis mental states. I then suggest that the analogy Chudnoff draws between

[1] Joshua Kelsall is currently taking an MLitt in Philosophy at the University of St Andrews, finishing in 2017. He completed his BA in Philosophy at the University of Leeds in 2016. His current research interests are history of philosophy, ethics, and political philosophy.

perceptions and intuitions does not hold, because perceptions are belief independent, while intuitions are not. I then respond to Chudnoff's counter-example to the belief view, which consists in the claim that it is possible to have an intuition, but no corresponding belief or inclination to believe. I will do this by providing him with an error theory that I hope, demonstrates that his counter-example to the belief view does not hold. By casting doubt on Chudnoff's notion of an intuition as sui generis mental state, and by responding to his criticisms against the belief view, I hope to vindicate the view that intuitions are inclinations to believe, which take our background beliefs and theories as input.

2 Kinds of Intuitions

Before we can begin to offer an analysis, we need to establish more precisely what constitutes an intuition, and ask what kinds of intuitions are relevant to this essay. In ordinary usage, there are lots of senses in which one might be said to have an intuition. For instance, consider a mother's intuition that her child is unwell, or the intuition of a lost traveller on which is the best way to find safety. The mother's intuition might be the result of maternal instinct, while the traveller's intuition may be based in an inexplicably good sense of direction. While such intuitions are irrelevant to this essay, it is important to highlight a common feature between all intuitions, including these ones. This common feature is that intuitions are all experiences in which it seems to the agent who has the intuition, that something is the case. For the mother, it seems to her that her child is unwell, and for the traveller, *it seems* to them that north is the best direction to go. All intuitions appear to have this quality of seeming, although often these seemings consist in different things, be it a maternal instinct, sense of direction etc. In this paper however, I will be concerned only with those intuitions employed and generated within philosophical argument and thought. In explaining such intuitions, consider the case where someone asks you to consider the proposition that "a fully rational person does not believe both P and not P" (Pust, 2016). The result of considering this proposition may be that it seems to you that it is true that a fully rational person does not believe both P and not P. These are the kinds of intuitions that a philosopher might use when thinking or arguing. Further examples include:

(1) It is impossible for a circle to be completely blue and completely yellow
(2) Murder is wrong
(3) S is only morally responsible for x, if S could have done otherwise

When I reflect on the proposition (2) it seems to me that it is the case that murder is wrong. When I reflect on proposition (3), it seems to me that it is true that S is only morally responsible for x, if S could have done otherwise. The same seeming occurs when considering (1). These philosophical intuitions are cases where intuitions emerge in virtue of our reflection on these propositions. The thought here is that there is something about these propositions, which results in our having intuitions about their truth (or falsity). However, this doesn't really tell us anything about what these intuitions are, or where they come from. One way of explaining where intuitions come from is to say that they are just beliefs that I have, and this is the view that I consider in the next section.

3 Intuitions as Beliefs

In this section I introduce the view that intuitions are beliefs. The three main views in this school of thought are as follows: 1) the *simple view* that intuitions are just like our ordinary beliefs, 2) the *aetiological view* that intuitions are beliefs with different aetiologies, and 3) the *inclination view* that intuitions are inclinations to believe (inclination view).

One account of the simple view comes from Lewis, who claims that "our "intuitions" are simply opinions; our philosophical theories are the same" (Lewis, 1983, x). Under this view, there is no significant difference between our intuitions and our ordinary beliefs. The main claim of this view is as follows:

(S) "S has the intuition that P, if and only if S believes that P" (Pust, 2016)

The aetiological view is the view that intuitions are beliefs with particular aetiologies. A good reason to hold this view stems from the thought that intuitions have characteristics that distinguish them from our ordinary beliefs, such as the spontaneous way in which they strike us, or their inexplicability. Gopnik and Schwitzgebel's (1998, 77) account of these features is that the aetiology of intuitions consists in subconscious processes, which we cannot directly observe. We can define this as:

(A) If S has the intuition that p, then S forms the belief as part of an unconscious reasoning process that S cannot directly observe.

The final belief view I will discuss is the view that intuitions are not beliefs, but inclinations to believe. Earlenbaugh and Molyneux defend this view, suggesting that intuitions are only a sub-class of inclinations to believe, which means that not every inclination to believe is an intuition (Earlenbaugh and Molyneux, 2009). We can characterise this view in the following way:

(I) If S has the intuition that p, then S has an inclination to believe that P

Of these three views, the aetiological view and inclination view are preferable to the simple view. The problem with the simple view, as noted in my introduction of the aetiological view, is that it doesn't really capture the unique flavour of intuition experiences; their spontaneity and inexplicability. For these reasons, I will not attempt to defend the simple belief view here, so from here on out my discussion will be entirely focused on the aetiological view and the inclination view.

4 Against Intuitions as Beliefs

In his paper, "What Are Intuitions Like" Elijah Chudnoff (2011, 626) argues against the belief, or to use his terminology, "doxastic" view, by providing two objections which he believes applies to any interpretation of the belief view. Each objection targets one of the following two general principles which Chudnoff takes belief views to be committed to:

"(DoxI1) Necessarily: If x has an intuition that p, then x judges, or has an inclination to judge, that p" (Chudnoff, 2011, 627)

"(DoxI2) Necessarily: If x has judges, or has an inclination to judge that p, then x has an intuition that p" (Chudnoff, 2011, 627)

It should be noted that Chudnoff uses the word "judgement" instead of "belief." The reason being that Chudnoff is arguing against doxastic accounts of intuition in general, so he wants a general term which captures all the different kinds of doxastic attitudes or judgements one might have regarding a proposition. Beliefs are themselves one kind of doxastic attitude or judgement, so they are captured by the term 'judgement' in (DoxI1) and (DoxI2).

Another thing to note is that belief views need not be committed to both (DoxI1) and (DoxI2). Chudnoff himself suggests that most doxasticists wouldn't commit themselves to (DoxI2) (Chudnoff, 2011, 627). I am inclined to agree that none of the belief views I wish to discuss are committed to (DoxI2).[2] In fact, Earlenbaugh and Molyneux, explicitly refute (DoxI2), by stating that not every inclination to believe amounts to an intuition (Earlenbaugh and Molyneux, 89). Nevertheless, the belief views discussed here are all committed to (DoxI1), since each view identifies intuitions with either judgements or inclinations to judge[3]. It follows from all three views then, that at a minimum, if one has an intuition that p, then one will have either a belief, aetiological belief, or an inclination to believe, depending on which of the belief views one ascribes to. For the simple view, any old belief is an intuition, because intuitions are just ordinary beliefs. For the aetiological view, intuitions will be a belief with a specific origin, which distinguishes them from our ordinary beliefs. For the inclination view, intuitions will just be a certain leaning, or willingness that we have to believe in whatever it is we are having and intuition about. Given that each view is committed to (DoxI1), if Chudnoff's objection against (DoxI1) holds, then it seems that it will hold against all three views.

Chudnoff's argument against (DoxI1) demonstrates that it is possible to have an intuition that p, without the belief or the inclination to believe that p. In other words, to show:

> "Not-(DoxI1): Possibly: x has an intuition that p and does not judge, or have any inclination to judge, that p" (Chudnoff, 2011, 632)

Chudnoff's begins his strategy for demonstrating that Not-(DoxI1) by drawing an analogy between intuitions and perceptions.[4] Chudnoff supports the sui generis view of intuitions, and the sui generis mental state that intuitions consist in is a unique experience, that can't be entirely explained in terms of our other mental states, and which is supposed to be similar to a perceptual experience[5]. For Chudnoff, our sense perceptions present the external world as being a certain way, while our intuitions present abstract matters to us in a certain way (Chudnoff, 2011, 626). An important similarity between these different experiences is that one can have an intuition that p, yet not believe, (or be inclined to believe) that p, in the same sense that one can perceive p, and yet not believe, (or be inclined to believe), that p.

First, let's unpick what it means for someone to perceive p, yet not believe that p? Consider the illusion of a hollow mask. In this illusion, the concave side of the

[2] I think that the simple intuition view of intuitions as opinions may be committed to (DoxI2), because the view does not make any distinction at all between intuitions and our other beliefs. However, given that my defence of the belief view focuses on the aetiological and inclination views, I need not discuss this further.

[3] Where we understand belief to be a kind of judgement.

[4] Chudnoff's arguments here take inspiration from Bealer (1998), who also believes that intuitions are akin to perceptions, and presents similar arguments.

[5] This isn't to say that intuitions are completely detached from our other mental states. Chudnoff stresses that in fact, our mental states such as beliefs, *constitute* our intuition experiences, in the same way that a lump of clay constitutes a statue. His idea then, is that our existing mental states can come together in such a way that they constitute a new intuition experience, just as the lump of clay, formed in a particular way, constitutes a new object, the statue (Chudnoff , 2011, 645)

mask visually appears to be the convex side of the mask. Once I am aware of the illusion, my perception of the concave side remains unaffected and I still perceive it as the convex side. However, I neither believe, nor am I inclined to believe that I am perceiving the convex side, because I know that I am perceiving an illusion. What I do have however, is a "sensory seeming"[6] that I am seeing the convex side of the mask. Nevertheless, Chudnoff (2011, 628) suggests, because our perceptions are belief independent, they will not change even if we know or believe that we are perceiving illusions.

I have established what it means to have a perception that p, but no belief or inclination to believe that p. However, we still need to explain how it is possible to have the intuition that p, yet not believe that p, which is Chudnoff's view. Take the intuition criticised by Gettier (1963), that the Justified True Belief (JTB) theory of knowledge is the correct theory of knowledge. Gettier objected to the intuition by presenting a counter-example to it, which showed that it is possible to have a justified true belief which does not amount to knowledge. Having read Gettier, I no longer believe in the justified true belief theory of knowledge, nor am I inclined to believe it. Nevertheless, Chudnoff would suggest that it is still possible that I have the intuition that JTB is true, despite my not believing, nor being inclined to believe it. In order to highlight how this is possible, consider another case from Bealer (1998, 208), who claims that this phenomenon occurs to him when he considers the naïve comprehension axiom. He does not believe, nor is he inclined to believe in it, but he claims that it intellectually seems to him that the axiom is true, which is to say that he has the intuition that it is true. This intellectual seeming works similarly to the sensory seeming in the hollow mask illusion. Just as the sensory seeming represents an image of the external world, the intellectual seeming represents an image of the abstract world. However, because I know that this intellectual seeming is false, I do not believe, nor am I inclined to believe it, just as I have no belief or inclination to believe that my perception of the mask is not an illusion.

The belief views I have raised are committed to the claim that if one has an intuition that p, then one has either some kind of belief that p, or an inclination to believe that p. Chudnoff and Bealer have provided counter-examples to that claim, whereby one still has an intuition that p, yet one has no belief, nor any inclination to believe that p. The challenge for a supporter of these belief views then, as Chudnoff (2011, 633) argues, is to provide an error theory that shows how Chudnoff and Bealer are going wrong when they say that they have an intuition without any corresponding belief or inclination to believe.

5 The Disanalogy Between Intuitions and Perceptions

Before discussing a potential error theory for Chudnoff and Bealer's counter-examples to the belief view, I would first like to discuss another objection to their theory, which focuses on Chudnoff's claim that intuitions are similar to perceptual experiences, and since beliefs are not perceptual, they cannot fully capture the nature of intuitions.[7] I will argue that in fact, intuitions have more in common with beliefs than intuitions. I argue that perceptions are belief independent, which is to say that perceptions do not change when our beliefs change. I argue that this independency does not occur with purported intellectual seemings, because

[6] This term comes from Bealer, "Intuition and the Autonomy of Philosophy" p.208

[7] I focus on this objection first, because it raises some important points which will become useful for the error theory that I discuss in section 6.

intuitions are belief dependent. I then suggest that the best explanation for the belief dependency of intuitions is not that they are intellectual seemings, presenting abstract matters to us in some way, but rather, they are inclinations to believe in a proposition. I will then argue that these inclinations result from an unconscious reasoning process which takes our current tacit and background beliefs and theories as their inputs.

A good reason to accept Chudnoff's account of intuitions as sui generis mental states is grounded in the fact that his theory accounts for the supposed similarities between intellectual seemings (intuitions) and sensory seemings (perceptions). Intuitions are sui generis mental states that present abstract matters to us, in the same way that sensory experiences present the external world to us. Chudnoff's objection in the previous section drew the analogy between perceiving a visual illusion and perceiving abstract phenomena such as naive comprehension. My objection to this analogy is to point out that our perceptions of illusions are belief independent,[8] while our intuitions appear to be belief-dependent. The thought being that our intuitions about a given phenomena may change as our beliefs about the world change, unlike our sensory perceptions of illusions, which remain constant even when our beliefs change. But if intuitions are supposed to be like perceptual experiences, then it seems as though they shouldn't be belief dependent, and should be consistent in the same way that sensory perceptions are consistent, which is to say that they don't change just because our beliefs change.

In the hollow mask illusion case, even when I know that I am seeing an illusion, this knowledge cannot alter my perception such that I no longer perceive the reality and not the illusion. However, when it comes to intuitions, they often change when my beliefs about the world change. In fact, often our intuitions disappear entirely over time as our beliefs change, unlike our sensory perceptions of illusions. Gopnik and Schwitzgebel (1998, 75-91) provide an example of this disappearance of intuition within developmental psychology. Their case study describes the differences between the intuitions of young children and fully-grown adults. One case focuses on the intuition that all beliefs are true, possessed by most three-year-olds. As children develop and learn more about the world, this intuition disappears, perhaps as children realise through experience that beliefs do not always correspond to truth. This causes a change in their understanding, and they no longer have the intuition that all beliefs are true. If we take for granted that most children have these intuitions, and if we suppose these intuitions are intellectual seemings in the way proposed by Chudnoff and Bealer, it seems strange that we no longer have access to these intellectual seemings now, when we consider the proposition of whether beliefs are always true. The fact that we don't believe it, or that we are not inclined to believe it shouldn't matter, because if it did matter, then that would suggest that intuitions consisted in our beliefs or inclinations to believe, but Bealer and Chudnoff want to insist that they are not.

Gopnik and Schwitzgebel's case study demonstrates how intuitions are affected by beliefs, because the intuitions of these children change when they develop and learn more about the world.[9] In other words, it demonstrates the belief dependency of intuitions. Moreover, these cases aren't limited to developmental psychology, they also occur within philosophy. Philosophers may (and often do) begin with intuitions that some theory is true, which disappear when they determine its falsehood. I for

[8] It is worth pointing out that Chudnoff, (quoting Evans) acknowledges this point. See Chudnoff, "What Intuitions Are Like,"p.628 and Gareth Evans, *The Varieties of Reference*, (Oxford University Press, 1982).

[9] This in fact is the preferred explanation of the developmental psychologists, of why the intuitions of young children differ from older children and adults. Ibid, p.91

instance, used to find the JTB theory intuitively true, but after reading and reflecting on Gettier counter-examples, I don't have any such intuition anymore. If intuitions are perceptual experiences, they should not be belief dependent in this way. When it comes to my sense perception of the illusion, I might know every last detail about the illusion, why it occurs etc., but none of this knowledge removes the sensory seeming. However, when it comes to supposed intellectual seemings, they seem to change depending on my beliefs, as they do in the case where I reflect on Gettier counter-examples to the JTB theory of knowledge. If this is the case, then it isn't clear that there is any good reason to suppose that intuitions are anything over and above beliefs or inclinations to believe, that emerge from our current understanding and beliefs about the world.

Chudnoff and Bealer could respond here by drawing a new analogy between perceptions and intuitions. Perhaps these belief dependent intuitions are not akin to perceptual experiences of illusions, but cases of perceiving things unclearly. Suppose I quickly glance out of my window and think I see a bird, but it turns out to be a balloon when I look more closely. My perception was mistaken because I wasn't looking clearly, and when I did look clearly, I saw the object for what it was. After seeing things clearly, I would not have the perception of a bird again because I am focusing clearly, thus my perception changes accordingly when I focus properly. In the same way, in the case of belief dependent intuitions, perhaps I'm just not focused enough to intellectually see things clearly enough. This seems apt for explaining why my childhood intuitions about beliefs have disappeared, because when I was a child I did not have much of a grasp on reality I do now, in other words, I wasn't seeing things clearly then as I am now.

In response to this, it is worth pointing out that it seems as though *all* intuitions are belief dependent. Even regarding Bealer's intuition about naïve comprehension, it is perfectly possible that somebody who used to have such an intuition, loses their intuition entirely once their understanding of the world changes enough.[10] If this is right, then I suggest that there aren't any intuitions which have the kind belief independency that perception has. My perceptions don't change depending my beliefs, yet my intuitions do often change depending on my beliefs. While Chudnoff and Bealer can account for this by drawing the analogy between intuitions and not perceiving things correctly, I think it is equally plausible to suggest that intuitions are just beliefs or inclinations to believe that result from our current understanding of the world, since all cases of intuition, I have suggested, seem belief dependent anyway. Given this, the belief view is no less plausible than the sui generis view. Moreover, the belief view is the more parsimonious of the two, as it doesn't posit any new entities such as sui generis mental states, and given that both views equally account for the same phenomena, we have good reason to favour the belief view.

6 Rejecting and Eliminating Intuitions

In the previous section I objected to Bealer and Chudnoff's claim that intuitions are similar to perceptual experiences. However, this hasn't done much to respond to Bealer's claim that when he comprehends the naïve comprehension axiom, he has an intuition but no belief nor any inclination to believe it. The task of this section is

[10] What exactly it means for a theory of the world to "change enough" in order to eliminate intuitions, will be discussed in the next section.

to explain what is occurring here, and why Bealer is wrong to say that he has no belief, or inclination to believe, in naïve comprehension.

I will go about this by defending Williamson's claim that what occurs in cases like Bealer's, is that he has the intuition that Naïve comprehension is true, which is an inclination to believe, which is then resisted because he knows better (Williamson, 2007, 217). Although this means adopting a defence of the inclination view in favour of the aetiological view, I will not be dismissing the aetiological view entirely for two reasons. First, I think Gopnik and Schwitzgebel are right to claim that intuitions result from unconscious processes driven by background theory. Moreover, I think it is perfectly reasonable to suppose that these unconscious processes driven by background do result in beliefs. My intuition that murder is wrong is the result of an unconscious process, nevertheless I also believe it.[11] The problem however, is that this isn't necessarily always the case, and Bealer's naïve comprehension example is one such case where there is no belief. What there is in this case, is an inclination to believe in naïve comprehension, which results from an unconscious process, and it is this inclination to believe that Bealer is missing.

Suppose it is true that intuitions are driven by unconscious or underground processes which cannot be directly observed, and that sometimes this process leads us to form beliefs. Contrast this process with a conscious thought processes which can also lead us to form beliefs. For beliefs that arise consciously, we are usually able reflect on the process which led to the belief, something which does not occur with intuitions driven by unconscious processes. What we have then, are two different processes, one unconsciously driven by background theory, and a conscious process driven primarily by our current conscious considerations.

I think that when Bealer has the intuition that naïve comprehension is true, there is a conflict between the unconscious and conscious processes. Bealer's intuition is an inclination to believe that naïve comprehension is false, which is caused by an unconscious reasoning process driven by background theory, which he cannot consciously observe. This intuition conflicts with Bealer's conscious thought process, which has led him to the conclusion that naïve comprehension is false. When Bealer is made aware of this conflict, or indeed, when anyone realises that they hold an inconsistent set of beliefs, he rejects his intuition.[12] However, I do not think that merely rejecting an inclination to believe will eliminate the inclination altogether. Bealer's error then, is to suppose that rejecting his intuition also eliminates the intuition.

The reason why the rejection of an intuition does not eliminate the intuition is due to the fact that intuitions, conceived as the fruits of unconscious processes driven by background theory, are not something that we have direct control over. I claim that the only way to eliminate an intuition is to change the background theory which the unconscious process takes as its input in generating intuitions. To see this point, think back to the children's intuitions about beliefs. In that case, it was suggested

[11] This isn't to say that the belief that murder is wrong, could not be the result of some kind of conscious reasoning process, it certainly could, if it was a conclusion that I came to through conscious reasoning. It is just to say that, if I have this belief, having not gone through any such conscious process, then this belief is just an intuition resulting from an unconscious reasoning process.

[12] Moreover, I think this rejection of the intuition is his reason for claiming that he has no belief, or inclination to believe. If he accepted the intuition, it seems he would claim to believe it, or if he suspended judgement, it would seem that he might claim to have an inclination to believe it.

that the false intuitions of the children only disappear once the child's background theory changes, which occurs as they gain a better understanding of the world around them (Gopnik and Schwitzgebel, 1998, 91). I suggest that in Bealer's case, his background theory has not changed enough in order to get rid of his intuition, and that is why he still has it.

Nevertheless, Bealer will still insist that he has no inclination to believe in naïve comprehension, and my considerations have not yet shown how he misses such an inclination. I think part of the solution to this, strange as it may seem, is that in a sense Bealer is right. He does have no inclination to accept naïve comprehension, in the sense that he has no inclination to accept naïve comprehension *as a part of his conscious reasoning processes.* Nevertheless, the inclination still emerges from his intuitive unconscious process. Bealer's error then, is to suppose that just because he is not inclined to accept naïve comprehension in his conscious account of reality, that he has no inclination *whatsoever,* to accept it. My response is that he does have such an inclination, an inclination resulting from this background process, which cannot be eliminated unless Bealer's background theory itself changes. A useful analogy here is that of a scientist who accepts counter-intuitive theories in their scientific image of the world because their theory demands it, even though their ordinary intuitions are inconsistent with it. Nevertheless, the scientist won't be inclined to using those intuitions in their scientific theories, because their scientific considerations won't allow for it, and this is also what occurs with Bealer. His background understanding of the world causes the inclination to believe that naïve comprehension is true, but Bealer is not inclined to accept this intuition into his explicit theory. Bealer takes this lack of inclination to accept the intuition as a part of his explicit account of reality as a lack of inclination altogether. I on the other hand, claim that Bealer cannot remove the inclination by simple refusing to admit it into his explicit theory; his background theory first must change, but then the intuition experience will go away altogether, thus, it is not possible to have an intuition that p, but no inclination *whatsoever,* to believe that p.

7 Conclusion

In this essay I argued that intuitions are inclinations to believe. The simple belief view failed to capture the spontaneity and inexplicability of intuitions, and while the aetiological view could account for these features, it faced Chudnoff and Bealer's counter-examples, whereby one has the intuition that p, yet no belief that p, regardless of the aetiological state. However, I argued that even Chudnoff and Bealer's view of intuitions as sui generis states fails, because their view is built on the claim that intuitions are like perceptions. I suggested that this view fails because intuitions are belief dependent, while perceptions are not. The idea of intuitions as inclinations to believe is the strongest view, I claim, because it offers the best account of the belief dependency of intuitions, as well as their spontaneity and inexplicability. Moreover, when construed as inclinations that result from unconscious processes which take our background theories and understanding of the world as input, we can explain Bealer's claim that he can have intuitions, yet no inclination to believe in those intuitions. The claim is that Bealer really does have no inclination to believe in those intuitions, as a part of his explicit theories. Nevertheless, until his background theory changes, those theories will still result in the inclination to believe in naïve comprehension, even if Bealer has no inclination to add it to his explicit theory.

References

Bealer, George, "Intuition and the Autonomy of Philosophy." in *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry,* edited by Michael R. DePaul and William Ramsey, 201-239. Rowan & Littlefield Publishers, 1998.

Chudnoff, Elijah. "What Intuitions Are Like." *Philosophy and Phenomenological Research* 82 (2011): 625-654.

Earlenbaugh, Joshua and Bernard Molyneux. "Intuitions are Inclinations to Believe." *Philosophical Studies* 145 (2009): 89-109.

Evans, Gareth. *The Varieties of Reference*, Oxford: Oxford University Press, 1982.

Gettier, Edmund. "Is Justified True Belief Knowledge?" *Analysis* 23 (1963): 121-123.

Gopnik, Alison., and Eric Schwitzgebel. "Whose Concepts Are They Anyway?" in *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry,* edited by Michael R. DePaul, and William Ramsey, 75-91. Rowan & Littlefield Publishers, 1998.

Lewis, David. *Philosophical Papers: Volume I*, New York: Oxford University Press, 1983.

Pust, Joel. "Intuition", *The Stanford Encyclopedia of Philosophy,* (Spring 2016 Edition), edited by Edward N. Zalta.

Williamson, Timothy. *The Philosophy of Philosophy*, Malden, MA: Blackwell, 2007.

A Parsimonious Agent Causation

Steven Guillemette[1]

Abstract

Free will and moral responsibility share an intricate relationship. Whether we deem an agent morally praiseworthy or morally blameworthy depends, in part, on whether or not the act was "performed freely". However, numerous difficulties emerge when one tries to articulate what "performed freely" entails: one must not only provide an account of free will with reference to its nature, he or she must also take into consideration its compatibility, or lack thereof, with existing and/or potential features of the world. Ultimately, this paper attempts to do just that, to provide an account of free will that is both coherent and naturalistic. Consequently, I argue that such an account be contrived under the umbrella of agent-causal libertarianism using the mechanics of quiescence.

1 Introduction

Whether we deem an agent as morally praiseworthy or morally blameworthy depends, in part, on whether the act was performed freely. However, numerous difficulties emerge when one tries to express what performed freely entails. It is my goal, then, to resolve these difficulties by providing a coherent account of free will.

In §1, I examine literature to supply an understanding of free will, including theories regarding its nature and its various debates. In §2, I consider contemporary objections to agent-causal libertarianism, a philosophical position concerning free will. In §3, I provide my own account of agent-causal libertarianism outlining its requirements in detail. Finally, in §4 I address possible objections to my account. Ultimately, it is my position that agent-causal libertarianism remains a tenable view worthy of consideration in current free will debates.

2 Free will and Determinism

There are two main conceptions of the nature of free will.[2] According to the first conception, an action is up to the agent if he "could have done otherwise," or, rather, if there were alternative possibilities open to him (Kane, 1998, 32). For example, one might say that the Greek hero Odysseus has free will in choosing to confront Scylla if he could have confronted Charybdis instead, or even refrained from entering the deadly strait altogether. According to the second, less familiar conception, an action is up to the agent if he is ultimately responsible for it (Kane, 1998, 35). Using the example above, one might say that Odysseus has free will if he, and nothing else, is the ultimate reason for his choice to confront Scylla. This holds, even if there were no other alternatives open to him at that time. Ordinarily, philosophers define and

[1] Steven Guillemette is a senior (graduating in 2017) at the University of St. Thomas. He majors in Philosophy and psychology and is interests in metaphysics and philosophical anthropology.

[2] Here, I am following the work of Kane (1998)

analyse free will along one of these two conceptions (van Inwagen, 1983); however, it is not uncommon to think free will involves both.

In addition to the nature of free will, much has been written regarding the thesis of causal determinism. In short, causal determinism assumes that all present and future occurrences are necessarily determined by past events and the laws of nature. Take, for example, the act of throwing dice. If one knew the position of the dice upon its release, its properties, and the laws of nature (i.e., the laws of gravitation, classical mechanics, etc.), then he could reasonably determine where the dice would come to rest after it was thrown (Timpe, 2013, 18). Now, according to the thesis of causal determinism, the world operates like the dice in the sense that the non-relational past and relevant laws of nature causally determine our behaviour.

Generally, philosophers fall into camps to explain how the thesis of causal determinism operates with respect to free will. Compatibilists argue that the existence of free will is compatible with the truth of causal determinism. Strictly speaking, if scientists were to discover that causal determinism is true, agents could still perform at least some actions freely. Incompatibilists, on the other hand, argue that the existence of free will is incompatible with the truth of causal determinism such that an agent acts freely only if casual determinism is false. In any case, it is important to point out, as Timpe suggests, that "neither position by itself is making a claim about whether or not agents actually do possess free will" (Timpe, 2013, 18). However, there are several subgroups which do make this claim, such as libertarianism.

3 Libertarianism

Libertarianism is an account on which incompatibilism is true. Thus, it holds that free will exists. Simply put, "one acts freely only if one's action was not determined – directly or indirectly – by forces outside one's control" (Moreland, 2009, 41). Consequently, libertarians bear the burden of proof in showing that the thesis of causal determinism is in some way false.[3] Assuming they accomplish this onerous endeavour (which, for the purposes of this paper, I will accept); libertarians still must show how their variety of freedom is possible and even intelligible.

3.1 Noncausation or Simple Indeterminism

Typically, there are three ways to explain libertarian freedom. One way, headed by Ginet (1990, 1997, and 2002) and McCann (1998, and 2012), is to assert "neither that a free action has any internal causal structure nor that it be caused by anything at all" (Clarke, 2003, 17). Accounts such as these are labelled non-causal or simple indeterministic because they do not require any positive conditions on freedom; even overt actions that are a function of basic decisions and reasons are noncausal. One consequence of these accounts, then, is that a decision can be completely up to the agent; however, as Clarke has pointed out, the problem with them is that they "fail to provide adequate accounts of active control and acting for reasons" (Clarke, 2003, 24). Assuming this objection holds, (and I think it does given Clarke's reasoning above,) one can find an intelligible account of libertarian free will only among causal views; that is, views that describe how an agent is able to cause her choices freely.

[3] For arguments, see Timpe (2013) and van Inwagen (1983).

3.2 Event-Causation

Event-causation is another way to explain libertarian freedom. On this view, both *relata* in the causal relation of a choice (i.e., the cause and effect) are taken to be events. However, unlike compatibilism, event-causal libertarianism maintains that some events are caused indeterministically. For example, suppose Ted is deliberating about whether to rock out with Bill tonight. However, he also wants to complete his final history report so he does not fail his history class, prompting his father to send him to military school in Alaska and dashing his dreams of forming a band. After carefully considering his reasons for performing each action, Ted decides to complete his history report. According to the event-causal libertarian, Ted's desire to complete his report causes his behaviour, but it does not sufficiently determine his behaviour – there is still a chance his alternative desire to rock out could have caused him to perform that action instead.

As with non-causal accounts, event-causal libertarians face several objections, such as a lack of rational explicability and loss of control.[4] If it is true that decisions are caused indeterministically, then it seems like the agent would lack sufficient control over his actions, thereby undermining his responsibility for that action. In response, some argue that the indeterminacy should be placed before the decision, in the deliberative process, rather than at the moment of decision (Mele, 1995). In this way, the agent can still be responsible for the action, even though it is caused indeterministically. However, as I explain below (§5.3), this response is problematic.

3.3 Agent-Causation

The final way to explain libertarian freedom and the account I adopt in this paper, is termed agent-causation. Agent-causation is similar to event-causation in that a free decision cannot be causally determined. However, it differs in that the cause of a free choice or volition (i.e., the first causal *relatum*) is not an event at all, but rather an agent. Randolph Clarke summarizes the requirements of such a view as follows:

> "On a common version of such a view, it is required (1) that a directly free decision or other directly free action be caused by the agent; (2) that neither the decision (or other action) nor the agent's causing that decision (or other action) be causally determined by events; and (3) that this causation by the agent not consist in causation by events."
>
> (Clarke, 2003, 133-134)

In such a way, the agent, or substance, is the ultimate source and sole originator of her free choices – she and nothing outside of her is responsible. Given this framework, one can examine a few problems facing this account.

4 Problems with Agent Causation

4.1 The Problem of Reason-Explanation

The first problem (or set of problems) relevant to agent-causal accounts concerns the reason-explanation of free choice. It is widely thought that if an agent acts according to reasons, then her recognition of those reasons, by citing them, will inevitably provide an explanation for why she chose to perform that action (Clarke, 2003, 21).

[4] For further objections to event-causation, see Clarke (2003).

In other words, reason-explanation contributes to an intelligible account of active control; if there are no reasons then there can be no active control.

One difficulty with reason-explanation is that agent-causal libertarians are committed to a non-reductive causal analysis; that is, they deny that events outside or within the agent (e.g. beliefs, dispositional states, reasons, etc.) can directly cause an agent's free choice. If they do concede this non-reductive analysis and argue that agential events directly cause a free choice, then they are no longer arguing for agent-causation, but rather event-causation. However, if the agent *does* remain committed to this analysis, it would seem that the agent has less control then if he did have reasons determining his behaviour. Ultimately, the agent-causal libertarian must provide an alternate account of reason-explanation to avoid overt mysteriousness and to justify active control.

4.2 Problem of Luck

Equally challenging is the problem of luck because it seems to impede agent's control over their actions. In the past several years, the "luck objection" has taken many different forms by philosophers both for and against agent-causal libertarianism; however, Balaguer's (2002) interpretation concisely captures the objection.[5] Simply put, "If a decision is undetermined at the moment of choice, then given the same past, the agent could just as easily have made another choice without anything about the agent changing, and so the agent could not be the source of, or have control over the choice" (Balaguer, 2002, 388).

Using the above example, suppose Ted istorn between rocking out with Bill and completing his final history report. Eventually, Ted chooses to complete his final history report in lieu of rocking out. Nevertheless, had one somehow turned back time, Ted could have easily chosen to rock out instead. It seems, then, that it is just a matter of luck, or perhaps chance, whichever choice Ted decides to make; and if it is just chance, then Ted is not really in control. Firstly, we should note that he problem of luck is not only a problem for libertarians, but also a problem for ethicists considering it is found in most philosophical discussions of moral responsibility. Nevertheless, it is still up to the agent-causal libertarian and not the ethicist to show that agential actions are not simply random since non-random, agential actions will contribute to a more coherent theory of agent causation.

5 A Naturalistic Account of Agent-Causation

The libertarian account defended in this paper is a modification of Rogers' account. It is rooted in Anselmian thought insofar as it employs a term coined by Anselm: *pervelle*, loosely translated as "per-will".[6] According to Rogers, the term "per-will…expresses the important technical concept that the choice for B [over A] is the *successful continuation of desiring B* to the end of the TC [torn condition] stage so it becomes the effective intention" (Rogers, 2015, 92). Given this description, one way an agent's free choice might proceed is as follows: at T(1) an agent, S, encounters a torn decision regarding whether to opt for A over B, two mutually exclusive events. Then, at T(2), following T(1), the torn decision terminates and the agent per-wills A

[5] For other arguments of this kind, Waller (1988), Mele (1999), Fischer (1999), Strawson (2000), and O'Connor (2000).

[6] See Anselm's De Casu Diaboli (Anselm and Hopkins, 2000) and Katherin Rogers (2015).

instead of B. However, immediately upon termination of the torn decision and thereafter, the agent could have per-willed B instead of A.

Initially, this account seems similar to the event-causal account described in §3.2 insofar as it gives prominence to torn decisions, a specific subset of decisions. However, it differs in a few crucial ways. First, as with all agent-causal accounts, the agent and not prior events, causes a directly free decision or action. Second, this account employs an Anselmian concept, *pervelle*, as a model for free choice based on quiescence. Finally, the relevant indeterminacy necessary for libertarian free choice occurs at the moment of choice rather than proceeding it.

5.1 Torn Decisions

According to Balaguer (2002, 382), "a torn decision is a decision in which the person in question (a) has reasons for two or more options and feels torn as to which set of reasons is strongest, i.e., has no conscious belief as to which option is best, given her reasons; and (b) decides without resolving this conflict."

Looking back at Ted's predicament, one might notice that certain reasons motivate his choice to rock out with Bill: certainly, it produces great pleasure for him, and helps him hone his skills as a guitarist for his band, "Wyld Stallyns". However, one might also notice that certain reasons motivate his choice to finish the history report: if he does not complete it, he will fail his history course prompting his father to send him to military school in Alaska and ending his career as musician. Suppose he deliberates for a while, but is unable to discern which set of reasons is stronger. Nevertheless, Ted chooses to complete his history report even though his reasons for doing so did not outweigh his reasons for rocking out. As a matter of course, he just chooses.

In many ways torn decisions are reminiscent of what Kane (1998, 107-115) calls "self-forming actions," which are "undetermined, regress stopping voluntary actions" that presuppose "plural rationality." However, there are two key differences. First, as Rogers (2015, 91) claims, torn decisions are mostly conscious experiences. It seems that agents are at least somewhat aware of their reasons for decisions, even though subconscious, dual willings (i.e. situations in which the agent has good reasons to choose A and B even though they are mutually exclusive) may come into play (Rogers, 2015, 91). Second, even though one may agree with Kane that if determinism is true, torn decisions could be causally determined by the natural universe (e.g., by events in our mind), one may disagree with Kane regarding where he locates the relevant indeterminism necessary for libertarian freedom. On his view, the indeterminism proceeds the choice, occurring during the torn condition, but as argued in §5.3, the indeterminism should be located at the moment of choice so as to avoid a further objection. At any rate, it is important to recognize that torn decisions involve a conscious, internal struggle between two mutually exclusive events, or motivating circumstances. Given this description, one may now turn to the Anselmian concept of *pervelle* to explain agent-causation.

5.2 Parsimonious Per-willing and Quiescence

On Rogers' account, the agent opts for B over A by per-willing (i.e. by continually desiring B until it becomes an effective intention) (Rogers, 2015, 96). Initially, it looks as if this account has a lot going for it: it does not introduce any new causal powers beyond the existing desires preceding the choice and the opting for B over A is under

the agent's control. However, it is still unclear how A becomes overridden the instant B is per-willed. If, at T(1), S desires both A and B, and then at T(2), he per-wills B so it becomes the intention, why does the desire for A simply go away? To some, this may seem oddly mysterious and, therefore, problematic. As a possible response to this objection, I propose the dropping out of a non-per-willed action be understood along the lines of quiescence, as explained below.

In Stump's *Aquinas*, she argues that the human will has three positions regarding some issue:

> "The will can assent to something or reject it, but it can also simply do nothing at all. It can just be turned off; it can be inactive or quiescent…Furthermore, in principle, the will can move directly from any one of these positions to another. That is, in general, it can move from rejecting to quiescence, from quiescence to assenting, from assenting to rejecting, and so on"
>
> (Stump, 2003, 394).

Given this description, Stump goes on to say that nothing external to the agent causes her state of quiescence, "it is always in the will's own power to be quiescent or not" (Stump, 2003, 401).

In order to explain how an agent's will becomes quiescent, Timpe (2007, 290) suggests we appeal to Dowe's analysis of "causation by omission". Simply put, Dowe's view is that causation by omission is not a genuine cause at all, but rather, "a counterfactual claim…about the mere possibility of causation" (Dowe, 2001, 216). In other words, omissions are cases in which a genuine cause is possible and therefore, "they can be cited in causal explanations. As a result, we can treat them as 'quasi-causes'." (Kittle, 2015, 92). Using this information as a framework, Timpe develops a theory of control whereby an agent controls an event *e* when either (1) an action of the agent causes *e* to occur; or (2) an omission by the agent quasi-causes *e* to occur (Timpe, 2007, 292). Now, applying the mechanics of quiescence and Timpe's theory of causation to Rogers' account of *pervelle* one can explain the dropping away of a non-per-willed action. To appreciate this, consider a free choice made by Ted. Suppose, at T(1), Ted encounters a torn decision regarding whether to rock out with Bill or to complete his final history report. Soon after, at T(2), Ted ceases to be torn and per-wills the completion of his history report until it becomes an effective intention. Immediately upon per-willing the completion of his history report, we can say that Ted becomes quiescent to his desire to rock out; that is, only when Ted per-wills the completion of his history report does Ted also quasi-cause the completion of his history report through the omission of rocking out. In this way, Ted is morally responsible for the outcomes of both his action and his non-action (i.e., his omission) in a manner that is not overtly mysterious.

5.3 Indeterminacy

Towards the end of §5.1 it was mentioned that placing the relevant indeterminism prior to the agent's choice could prove problematic for the libertarian. If, as Kane suggests, "[a free choice] one way or the other is *undetermined* because the process preceding it and potentially terminating in it (i.e., the effort of will to overcome temptation) is *indeterminate*," then the indeterminate process (viz., neurons firing one way or another) could probabilistically cause the agent's behaviour, introducing an element of luck (Kane, 1998, 128). For example, suppose Tom, through some

indeterminate process, deliberates between two mutually exclusive events, A and B. After serious thought, Tom ceases to deliberate and his desire for A probabilistically motivates his choice to A. However, it just may be that the indeterminate process preceding the choice terminates in neurons randomly firing one way when they could have fired another way. Ultimately, it does seem lucky that this indeterminate process could be responsible for causing an undetermined choice.

As a response to this objection, I suggest one place the relevant indeterminacy at the moment of choice. In doing so, external events (e.g., indeterminate processes, neurons, or the relevant laws of physics) could not cause the agent to choose as she does (Balaguer, 2002, 387). And if one juxtaposes a lack of external causation with intentional choice, as Balaguer suggests, then it is true to say that the agent is the sole originator of her free choices; that is, she has complete authorship and control over her free choices (ibid, 387).

6 Agent-Causal Problems Revisited

In §4, I introduced a couple problems that could threaten the adequacy of any agent-causal libertarian account. These problems are the reason-explanation of libertarian free choice and the problem of luck. The remainder of this paper dispels these threats with respect to the account provided above.

6.1. Problem of Reason-Explanation Revisited

Recall from §4.1 that the agent-causal libertarian is committed to a nonreductive analysis; that is, they deny that events outside or within the agent (e.g., beliefs, dispositional states, etc.) can directly cause a free choice. However, as the objection goes, if there are no reasons that cause an agent's actions or behaviour, then she has no active control over her actions or behaviour. Ultimately, the task of the agent-causal libertarian is to provide an alternate account of reason-explanation to justify active control.

Like Leibniz (1714) and Chrisholm (1964), I would argue that an agent's reasons motivate or incline without necessitating the making of a free decision, since necessitation would preclude agent causation. What this means is that, "[reasons] do not, on their own, cause the agent to decide as she does. Rather what makes the relevant difference is the agent's causal contribution to the decision" (Haji, 2008, 178). In such a way, an agent's free choice is motivated by a reasons-based element and caused by agent-causal element. To appreciate this, consider again a free choice made by Tom. Suppose, at T(1), Tom encounters a torn decision regarding whether to order a salad or a burger, two mutually exclusive events. Before long, at T(2), Tom ceases to be torn and per-wills that he order a salad in lieu of ordering a burger. The point here, is that Tom's reasons for ordering a salad or a burger at T(1) motivate his choice to order a salad or a burger, but they do not necessitate his choice. It is the inclusion of the agent-causal element that allows Tom to per-will, thereby causing, his ordering of the salad at T(2). Ultimately, this account of reasons-explanation grants the active control the agent-causal libertarian seeks without being arbitrary. However, it is still unclear why the agent chooses one option over the other, a problem to which I now turn.

6.2 Problem of Luck Revisited

As stated above, the problem of luck goes something like this:

> "If a decision is undetermined at the moment of choice, then given the same past, the agent could just as easily have made another choice without anything about the agent changing, and so the agent could not be the source of, or have control over the choice"
>
> (Balaguer, 2002, 389)

However, this objection does not hold. As Balaguer argues, even though Ted was torn at the moment of decision, meaning neither set of reasons pick out the best option, it does seem like he has a rationale for making a choice: namely, that he does not remain in a state of indecision (ibid, 389). Consequently, Ted has good reason to choose one option over the other, even though one option might not be clearly better than the other. Furthermore, we might say that what explains Ted's reasons for choosing to finish his history report in lieu of rocking out is simply Ted's acceptance of his reasons for that choice (Pruss, 2006, 135). That is to say, his reasons for completing his history report provide sufficient explanation for his choice even if they were not necessitating reasons. Had he chosen to rock out instead, his reasons for that decision would sufficiently explain his choice. Ultimately, though, it is important to understand that whichever choice Ted decides to make, he is still the source of his choice and in control insofar as his decision was conscious and nothing external caused him to choose as he did.

7 Conclusion

The purpose of this paper is to offer a naturalistic and coherent account of agent-causation that is not overtly mysterious or incoherent. Using Rogers' account as a starting point, I have argued that the dropping away of a non-per-willed desire is best understood using the mechanics of quiescence. In addition, I have demonstrated that the relevant indeterminacy necessary for libertarian free will should occur at the moment of choice as opposed to proceeding it. Suitably, the next step for agent-causal libertarians is to develop their understanding of the nature of indeterminacy, that is, to understand how it might occur in the process of free choice and in what proximity. Ultimately, such an understanding will have vast implications on our understanding of free will and moral responsibility.

Bibliography

Anselm, and Jasper Hopkins. *Complete Philosophical and Theological Treatises of Anselm of Canterbury*. Minneapolis: A.J. Banning Press, 2000.

Balaguer, Mark. "A Coherent, Naturalistic, and Plausible Formulation of Free Will." *Nous* 36 (2002): 379-406.

Chrisholm, Roderick. “Human Freedom and the Self.” in *Free Will*, edited by Robert Kane. Blackwell, 1964.

Clarke, Randolph. *Libertarian Accounts of Free Will*. Oxford: Oxford University Press, 2003.

Dowe, Phil. “A Counterfactual Theory of Prevention and ‘Causation’ by Omission.” *Australasian Journal of Philosophy* 79 (2001): 216-226.

Fischer, John M. “Recent Work on Moral Responsibility.” *Ethics* 110 (1999): 93-139

Ginet, Carl. “Freedom, Responsibility, and Agency.” *Journal of Ethics* 1 (1997): 85-98.

Ginet, Carl. *On Action*. Cambridge: Cambridge University Press, 1990.

Ginet, Carl. “Reason Explanations of Action: Causalist and Noncausalist Accounts.” In *The Oxford Handbook on Free Will*, edited by Robert H. Kane, 386-405. Oxford: Oxford University Press, 2002.

Haji, Ishtiyaque. *Incompatibilism’s Allure: Principal Arguments for Incompatibilism*. Ontario: Broadview Press, 2008.

Kane, Robert. *The Significance of Free Will*. New York: Oxford University Press, 1998.

Kittle, Simon. “Grace and Free Will: Quiescence and Control.” *Journal of Analytic Theology* 3 (2015): 89-108.

Leibniz, Gottfried. “Monadology.” In *Leibniz Selections*, edited by Philip P. Wiener. New York: Charles Scribner’s Sons, 1951 [1714].

McCann, Hugh J. *The Works of Agency: On Human Action, Will, and Freedom*. Ithaca: Cornell University Press, 1998.

McCann, Hugh J. “Making Decision.” *Philosophical Issues* 22 (2012): 246-263.

Mele, Alfred. “Kane, Luck, and the Significance of Free Will.” *Philosophical Explorations* 2 (1999): 96-104.

Mele, Alfred. *Autonomous Agents: From Self-Control to Autonomy*. New York: Oxford University Press, 1995.

Moreland, James P. *The Recalcitrant Imago Dei: Human Persons and the Failure of Naturalism*. London: SCM in association with the Center of Theology and Philosophy, University of Nottingham, 2009.

O’Connor, Timothy. *Persons and Causes*. New York: Oxford University Press, 2000.

O’Connor, Timothy. “Agent Causation.” In *Agents, Causes, and Events: Essays on Indeterminism and Free Will*, edited by Timothy O’Connor, 173-200. New York: Oxford University Press, 1995.

Pruss, Alexander. *The Principle of Sufficient Reason: A Reassessment*. New York: Cambridge University Press, 2006.

Rogers, Katherin A. *Freedom and Self-Creation: Anselmian Libertarianism*. Oxford: Oxford University Press, 2015.

Strawson, George. "The Unhelpfulness of Indeterminism." *Philosophy and Phenomenological Research* 60 (2000): 149-150.

Stump, Eleaonore. *Aquinas*. London: Routledge, 2003.

Timpe, Kevin. "Grace and Controlling What We Do Not Cause." *Faith and Philosophy* 24 (2007): 284-299.

Timpe, Kevin. *Free Will: Sourcehood and its Alternatives*. New York: Bloomsbury Press, 2013.

van Inwagen, Peter. *An Essay on Free Will*. Oxford: Clarendon Press, 1983.

Waller, Bruce. "Free Will Gone Out of Control: A Review of R. Kane's *Free Will and Values*." *Behaviorism* 16 (1988): 149-162.

Does Rousseau's discussion of sovereignty and government suggest a preference for elite domination?

Dominic Dark[1]

Abstract

The aim of this essay is to defend the thesis that Rousseau favoured popular sovereignty. Rousseau defines popular sovereignty as the "exercise of the general will" of the people, where the general will is those interests shared by the whole citizenry (Rousseau 2011, 170) . Before reaching this conclusion, I initially draw a distinction between the two ways in which Rousseau describes the function of government. On the one hand, it is described as the executive agent of the citizenry (which I call description A). On the other hand, he describes government as an independent body that may choose to prioritise its own interests (description B). Through the essay I argue that these two descriptions are incompatible, and that Rousseau's conflation of these two descriptions has engendered confusion as to whether he did in fact support popular sovereignty. The purpose of this essay is to defend the view that Rousseau did favour popular sovereignty, despite the extensive power that he describes the government as having.

1 Introduction

This essay draws a distinction between the two ways in which Rousseau describes the function of government. On the one hand, it is described as the executive agent of the citizenry. On this view the government just administers the laws that have been made by the people, applying it to particular instances and ensuring that it is upheld. By this description, (which I call description A) the government is in the service of the people. On the other hand, he describes government as an independent body that may choose to prioritise its own interests (description B). On this view, the government should apply the law as dictated by the people but has the capacity to impose, on the people as subjects, laws that pursue its own interests. Broadly speaking, critics have seen A as what Rousseau thinks the government ought to be like. Theoretically, the government should act only on the will of the people to execute laws. By contrast, critics have interpreted B as the role Rousseau thinks the government would have to hold in practice in order to govern effectively. As such, it is not clear whether Rousseau supported popular sovereignty, where the people are the ultimate source of authority, or a strong and potentially oppressive government. The purpose of this essay is to defend the view that Rousseau did favour popular sovereignty, despite the extensive power the he describes the government as having. However, to reach this conclusion we have to reconsider the purpose of Rousseau's sharp distinction between the government as the executive authority and the people as legislative power in description A.

In §2, I introduce Rousseau's distinction between the government and sovereign bodies. Through drawing on his concept of will, I explain Rousseau's motivation to

[1] Dominic Dark is studying Economics and Philosophy at the University of St Andrews and will graduate in June 2017. He has a broad range of interests in Philosophy, however is currently especially focused on functionalist accounts of the mind-body relation

separate executive power from legislative power into distinct political bodies. The section concludes with the question of whether Rousseau is less concerned about the prospect of members of government subverting the general will than the people. In §3, I substantiate the view that Rousseau was more concerned with ensuring efficient governance than the potential threat of a strong government undermining the general will. This conclusion draws heavily on the inference that Rousseau advocated a self-perpetuating government – one that elects its own future members. Such a result, if true, would appear to be highly undemocratic. In the final section, I show that the inference that Rousseau favoured a self-perpetuating government is based on an incorrect reading of the text. I argue instead that the people are intermittently given the executive power to elect government officials, as one would expect of a democratic theorist.

2 The relationship between sovereignty and government

In this section I distinguish between Rousseau's two political bodies, the government and the sovereign people. I explain that whilst legislative power, the ability to make laws, is held by the sovereign people, executive power, the ability to interpret and administer the laws, is often best entrusted to a government as a distinct body.

Just as the actions of a human body have two causes, so do the actions of the body politic. In order for a human body to perform an action such as jumping, it is necessary that there is both a) the will to jump and b) the physical capacity of the body to jump. Similarly, there are two conditions jointly necessary for the body politic to perform an action. On one hand, there must be a will to perform an action, the general will of the people expressed through legislation, and force given to that will, the executive power of the government (Rousseau 2011, 194). The general will is general in two senses. On one hand, it is general in the sense that each vote matters equally and on another, it is general in that it applies indiscriminately to all (Rousseau 2011, 172) . Therefore, we should not think of legislation as detailed legal documents, but as an expression of what the people collectively want. Executive power deals with the specific application of law.

This analogy describes the different roles of the government and sovereign, but does not explain why there needs to be two different bodies. It is possible that the people could have legislative and executive power. However, Rousseau is sceptical about the people's ability to govern itself well. A good government gives force to the laws, as expressions of the general will. An expression of the general will occurs when each citizen votes in favour of what they think is in the shared interests of the citizenry as a whole, the common good (Rousseau 2011, 172) . Were the citizenry to have the power to decide how a law was executed, in cases where their personal interests were involved, they would be liable to vote in favour of their private will ahead of their judgement of the common good. Imagine a self-governing people faced with the decision of whether or not to increase the national debt. Whilst it might not be in the common interest to raise the public debt level, if a citizen also had the power to decide how to spend the capital raised (by investing solely in their locality) they still might vote in favour of the increase, using it to further their personal interests. When each citizen puts their personal interests ahead of their judgement of the common interests, the sum of their private wills differs from the general will. This problem motivates Rousseau's clear distinction between the legislative sovereign and executive government, as in the human body analogy (Rousseau 2011, 197) (Bertram 2004, 150).

However, in practice, members of the government (*magistrates*) are just as susceptible to the temptation of using their executive power to gratify personal interests, or as a group, their collective interests, known as the corporate will.

In a perfect scenario, the *magistrate*s would act on the general will above both the corporate will and their private will as individuals. However, Rousseau argues that "the various wills become more active in proportion as they are more concentrated" (Rousseau 2011, 196) . I interpret this as meaning that each *magistrate* is inclined to pursue their private interests firstly, above the interests of the *prince* (the group of *magistrates*), and in turn, above the common interests of the citizenry. Whilst this does not mean *magistrate*s inevitably always pursue the corporate will over the general will, at the very least, Rousseau thinks this is a persistent worry.

Through this section I have drawn a distinction between the sovereign and the government, and have explained Rousseau's motivation for urging that the executive powers ought not be held by the people. However, noting that the government is just as prone to pursuing their own interest prompts the question of whether Rousseau is less concerned about their subverting of the general will. I address this concern in the following section.

3 How democratic are Rousseau's institutions?

Although Rousseau describes the role of the government as a servant of the sovereign people, critics have argued that the actual relationship between the two institutions is one of domination by the government (Bertram 2004, 173). Whilst Rousseau *claimed* to favour popular sovereignty, this objection holds that there is a tension between his purported republican aims and the restrictions he placed on the power of the people (Goldschmidt 1980, 153) (Fralin 1978, 90). I substantiate this claim in this part of the essay.

Rousseau thinks that it is inevitable that all states eventually collapse. A state collapses when the sovereign people is oppressed by the government. An oppressive government is one that acts only on the will of the *magistrate*s and ignores the general will of the people. This is an inevitable outcome of any state, since the government's corporate will "makes a continual effort against sovereignty" (Rousseau 2011, 212). I take the "continual effort" of the *magistrate*s to be a result of their stronger inclination to act on the more concentrated corporate will over the general will, as explained in §1.

However, Rousseau urges that whilst all states eventually collapse, some states last longer than others. A state's longevity depends upon the relationship between the sovereign people and the government, such that the likelihood of the government subverting the will of the sovereign is diminished. When there is little discrepancy between the private will and the general will, direct democracy, as a form of government in which the sovereign people have both legislative and executive power, can in fact flourish according to Rousseau. Though for a direct democracy to be stable, it must both be small and roughly equal in terms of wealth, such that citizens can assemble regularly to foster cultural bonds, sharing similar values and aspirations. As a third necessary condition, the inhabitants of a direct democracy must not enjoy luxuries. Rousseau argues that luxury causes citizens to become selfish and vain, two qualities conducive to prioritising personal interests above the common good (Rousseau 2011, 199). Under these three conditions, Rousseau thinks that each citizen more easily grasps the interests common to the community and,

presumably as a result of this, is less inclined to prioritise their private will over the general will (Bertram 2004, 157).

However, Rousseau thinks these necessary conditions for direct democracy are exceptionally demanding on citizens. In particular, his main concern appears to be that each citizen, having both legislative and executive power, is unlikely to prioritise the general will above their private will – as explained in §1. In writing about democracy, that "so perfect a government is not suited to men", Rousseau is suggesting that on the most part, humans would struggle to prioritise common interests over their personal interests were they to exercise executive power (Rousseau 2011, 200).

The implication of Rousseau's scepticism about the plausibility of direct democracy is that for most populations, in order for a state to be lasting, the executive government ought to be separate from the legislative sovereign. However, if we are to have a separate government, we return to the same problem from §1 – that the *magistrate*s are more inclined to pursue their personal interests over the general will.

Recall the distinction in the introduction between Rousseau's two descriptions of government, A and B. Fralin forcefully argues that these two descriptions are incompatible. Whilst it is the case theoretically, that the executive acts only on the will of the sovereign (A), in reality, as an intermediate political body that has the power to apply laws to subjects, the executive does have corporate will which it has the power to act on (B) (Fralin 1978, 92). Moreover, since the people cannot propose laws, but can only vote in favour or against government proposals, the government must bear significant legislative power in order to formulate these proposals in the first place (Fralin 1978, 92). Since the government cannot both be a partly legislative body with a will of its own, and at the same time be a purely executive agent of the people without the ability to act on its own will, these two depictions of government (A and B) are incompatible.

In practice, by description B, the government as a powerful independent body of *magistrate*s does not act as an agent but as a master of the sovereign people. For Fralin, Rousseau favoured dominant government power, not for any malicious reason but because a government needs extensive power, both legislative and executive, to govern effectively (Fralin 1978, 90).

It is important to note that whilst Rousseau explicitly thought elective aristocracy (rule by elected *magistrate*s) was often the most suitable form of government for a people, by itself this approval does not entail a preference for government domination. At least in theory, elective aristocratic government is completely compatible with popular sovereignty so long as the general will is not subverted by the corporate will of the government. Rather, it is through Rousseau's description of elective aristocracy as an intermediary body that Fralin argues citizens bear very little actual power.

Given the constraints of this essay, I want to focus on one of the most important restrictions on popular power that motivate this point of view. Fralin points to a passage which suggests that the government is self-perpetuating. The government alone decides the *magistrate*s which take office in the future (Fralin 1978, 108) (Rousseau 2011, 222). This makes sense, Fralin continues, since the executive's role is to deal in particular matters, and the choice of a *magistrate* is a specific matter.

That the government in Rousseau's elective aristocracy is self-perpetuating is taken as evidence, by Fralin, that Rousseau favoured powerful government rule that comes into conflict with his purported concern for popular sovereignty (Fralin 1978, 109).

As mentioned, this need not be interpreted as an artfully disguised inclination for cruel oligarchy. His insistence on strong government power can be thought of as a practical means of securing effective governance.

4 In defence of Rousseau's preference for popular sovereignty

In this final section I respond to the interpretation that Rousseau is advocating dominance of the people by a self-perpetuating elite government. I ultimately aim to show that Rousseau's notion of an elective aristocratic government is compatible with the democratic notion of popular sovereignty, provided we reconsider the strict distinction between the people as legislative power and the government as executive power suggested by description A of government.

Firstly, it is important to distinguish representative sovereignty, which Rousseau does not think is possible, and representative government, which Rousseau thinks is a practical necessity for most states. Sovereignty is defined by Rousseau as the "exercise of the general will", which I understand to mean the act of expressing the general will at lawful assemblies (Rousseau 2011, 170) . Since the general will is the interests that are common to the people, it follows that sovereignty is an expression of the interests common to people. Therefore, sovereignty is not the type of thing that can be represented, it has to come from the people by definition. On the other hand, representative *government* is the delegation of executive power to an independent body, which is both possible and encouraged, for the reasons given in §1. Fralin is not arguing that the government represents the sovereign. Rather that the sovereign power of the people is so limited and the legislative and executive powers of the government so strong, that the government inevitably acts on its own corporate will subverting the general will.

However, whilst it is inevitable in any state that the government eventually subverts the sovereignty of the people, it is possible that elective aristocracies are compatible with popular sovereignty, at least in the short term. Were this not the case, Rousseau would not bother with discussing how to place checks on the power of the government by legally guaranteed periodic assemblies (Rousseau 2011, 223) . He notes that "once the populace is legitimately assembled as a sovereign body, all jurisdiction of the government ceases" (Rousseau 2011, 217) . Once assembled, the government no longer exists and the people must ask itself two questions, "Does it please the sovereign to preserve the present form of government?" and "Does it please the people to leave its administration to those who are now in charge of it?" (Rousseau 2011, 224). These intermittent assemblies provide the sovereign with a means of overthrowing government and provide a clear incentive for *magistrate*s to act not on their corporate will but the general will of the people. That Rousseau advocates regular assemblies of the sovereign, during which time the government is liable to being overthrown, does not fit well with Fralin's picture of government domination. Rather than being master of the people, the ability to overthrow the government better fits the picture of the government being in the people's service.

One of the most important points that motivates Fralin's conclusion, that the government dominates the people, comes from his interpretation that the government chooses its own successor (Fralin 1978, 108) (Rousseau 2011, 222) . However, a close reading of the passage that motivates Fralin's reading, Chapter 17, Book 3 of the Social Contract, suggests that the people have this executive power, not the government. Once the people have assembled and decided that there "will be

a governing body", we are told that the nomination of leaders is "a function of the government" (Rousseau 2011, 222). This is a particular act, and particular acts are dealt with by the executive. However this creates a problem, since it is not clear how an executive government can decide who to nominate to leadership when it does not yet exist.

What immediately follows suggests that the people assume this executive power. The people "can in certain circumstances become *prince*" (Rousseau 2011, 222). Remembering that *prince* is the collective of *magistrates,* Rousseau is saying that the people on occasion take on the executive power usually held by the government. Cohen interprets this passage as saying that when no government exists, as is the case whenever the people are lawfully assembled, the executive authority to choose leaders belongs to the people (Cohen 2010, 169). At these points, there is a "sudden conversion of sovereignty into democracy" (Rousseau 2011, 222). Both legislative and executive power are combined in the people once assembled.

If it is the case that the people intermittently hold executive authority, when Rousseau talks about the nomination of leaders as being a "function of the government", he must be using "government" in a different way to when he strictly separates the legislative people from executive government (Cohen 2010, 170). Cohen makes sense of this by separating Rousseau's normative use of "government", what the function of government as executive ought to be, from his institutional use of "government", government in practice as the intermediate body standing between the sovereign and subjects (Cohen 2010, 170). In much the same way as Fralin, Cohen argues that whilst the government ought to be just the executive agent of the people (description A), in practice it is an intermediate body that has to assume some legislative authority in order to go about the day to day making of laws (description B). What Fralin failed to realise though, was that the sovereign people also holds significant power, by being able to vote in and out those officials in government (Cohen, 170).

5 Conclusion

Through this essay, I have defended the view that Rousseau favoured popular sovereignty rather than domination by an elite government. However, in reaching this conclusion I have urged that we ought not strictly separate the government as the executive force from the sovereign people as the legislative power. In practice, the government must exert some legislative power in making and proposing quotidian laws and the people intermittently hold executive power when voting in elections. Nonetheless, we should not just dismiss definition A of government, as the executive agent of the people. Importantly, this definition serves to emphasise that the government ought not undermine the general will (Bertram 2004, 174). Rousseau's description of government as both a powerful independent body and a subservient agent of the people reveals a practical concern to balance the need for effective government against executive transparency.

Bibliography

Bertram, Christopher. *Routledge Philosophy Guidebook to Rousseau and The Social Contract* . London: Routledge, 2004.

Cohen, Joshua. *Rousseau: a free community of equals*. Oxford: Oxford University Press, 2010.

Fralin, Richard. *Rousseau and Representation: a study of the development of his concept of political institutions*. New York: Columbia University Press, 1978.

Goldschmidt, Maure. Rousseau and Representation. *Political Theory 8* (1980), 551-554.

Rousseau, Jean-Jaques. "On the Social Contract, or Principles of Political Right." in *Jean-Jaques Rousseau The Basic Political Writings,* edited by Donald A. Cress and David Wootten, translated by Donald A. Cress, 153-252, Cambridge: Hackett Publishing Company Inc, 2011.