

Yale—UConn Graduate Student Philosophy Conference (2009):
Kantian Constructivism and Self-Legislation

*THIS PAPER IS A DRAFT. PLEASE DO NOT QUOTE, CITE, COPY, OR
CIRCULATE WITHOUT PERMISSION*

Kantian Constructivism and Self-Legislation

Nathaniel Jezzi
Cornell University
nj2@cornell.edu

Abstract:

John Rawls has argued that the constructivist's notion of objectivity provides us with all the objectivity we need to make sense of morality. That is, objectivity "understood in terms of a suitably constructed social point of view that all can accept"¹ coheres well with our moral concerns and beliefs. In this paper, I will not directly challenge the claim that the constructivist's objectivity is objective *enough*. Others have cast doubt on the idea that there is anything to be gained by such an attack.² Instead, I will assume for the sake of argument that this alternative notion of objectivity is plausible. The focus of my examination will be the constructivist's argument for the claim that her notion of objectivity coheres better with our moral concern for autonomy than the realist's does. Central to this argument is an interpretation of autonomy as requiring that we self-legislate the terms of morality. I argue that the *Self-Legislation* thesis does not withstand critical reflection. The traditional motivation for this thesis appeals to an argument that is either sound but uninteresting or interesting but unsound. Although Rawls introduces an innovative argument for *Self-Legislation* that appeals to our considered judgments about the fairness of decision procedures like his *Original Position*, this argument also fails. For, upon closer inspection, it turns out that the ends Rawls sets out for his procedure do not support an interpretation of autonomy as *Self-Legislation*. In fact, they look as if they speak against the thesis.

1. Autonomy and Kantian Constructivism

Kantian constructivism is the view that moral truths are determined by the choices that rational agents would make if they were to follow a procedure that allowed them to fully express their autonomy.³ It is a view that builds the demands of autonomy directly into the truth-conditions for all moral statements. Defenders of Kantian constructivism argue that the moral realist's notion of objectivity, by contrast, cannot sensibly be reconciled with the importance of autonomy. The realist posits a moral order that is independent of and prior to the judgments of moral agents. But constructivists claim that this is

Yale—UConn Graduate Student Philosophy Conference (2009):
Kantian Constructivism and Self-Legislation

incompatible with the concern that rational agents are self-governing. Although constructivists wield different arguments against moral realism, *The Argument from Autonomy* is arguably the only one that uniquely secures their doctrine as the preferred alternative to realism.⁴ It has two aims: first, it is an argument against moral realism; second, it lends positive support to Kantian constructivism.

The Argument from Autonomy

1. At least part of what it is to be a moral agent is to be self-governing or *autonomous*.
2. Autonomy requires that rational agents provide themselves with their own true moral principles. [*Self-Legislation*]
3. If *Self-Legislation* is true, then moral realism is false.
4. However, the demands of *Self-Legislation* are satisfied by Kantian constructivism.
5. So, Kantian constructivism bests moral realism.

This argument starts out with an appeal to a basic datum in moral phenomenology. It assumes that our thinking about moral obligation is constrained by a general belief that moral agents are autonomous. We think that part of what it is to be a moral agent is to be self-governing, and we accord agents a certain respect in virtue of this capacity. This belief forms a framework for determining what can count as an obligation, and it generates a concern to understand what our obligations are.

Since autonomy is such a concern, we must make space for autonomy in our moral theorizing.⁵ But, as it stands, we have not said enough about what *autonomy* is. Consequently, it is hard to know how it affects the nature of moral obligation and why this is something we should care about. At this point, the constructivist introduces her favored interpretation of autonomy as *living under one's own law*. However, there is an

Yale—UConn Graduate Student Philosophy Conference (2009):
Kantian Constructivism and Self-Legislation

obvious lacuna in this transition from premise (1) to premise (2). Why should we think that autonomy requires *Self-Legislation* [SL]?

There are two ways that constructivists have typically motivated SL. The first way involves an appeal to what I will call “The Kantian Argument”. It is inspired by an argument that Kant makes in *Groundwork, Section III* (446-8).⁶ Whether Kant himself would have accepted the constructivist’s interpretation of this argument is not what is at issue. In fact, there is a live dispute in the literature as to whether Kant himself is committed to the interpretation of SL that is taken as the conclusion of this argument.⁷ In any case, there are contemporary ethicists, such as John Rawls and Christine Korsgaard, who put forward something like this argument in support of SL, claiming Kant as their inspiration.⁸ The problem with this first argument, as I will show, is that it does not appear to secure the desired conclusion. For the argument is either sound but uninteresting or interesting but unsound. The second way of motivating SL appeals to our considered moral judgments about the appropriateness of Rawls’s famous deliberative procedure, the *Original Position*. Although this strategy of supporting the move from (1) to (2) appears more promising, it also leads to a dead-end for the constructivist. I will argue for an interpretation of Rawls’s *Original Position* that at best does not support constructivism and at worst is actually incompatible with the view.

2. One Traditional Argument for *Self-Legislation*

The following is a reconstruction of the Kant-inspired argument for SL. Those familiar with Kant’s own argument will note that this reconstruction omits both some terminology as well as some premises that are included in the traditional presentation.⁹

Yale—UConn Graduate Student Philosophy Conference (2009):
Kantian Constructivism and Self-Legislation

However, I would claim that it reasonably approximates what the traditional argument looks like when it is extracted from the Kantian apparatus.

The Kantian Argument

1. At least part of what it is to be a moral agent is to be self-governing or *autonomous*.
2. The moral law is either self-imposed or externally-imposed.
3. If the moral law were externally-imposed, rational agents could not count as acting on their own reasons when they act morally.
4. Acting on your own reasons is part of what it means to be self-governing.
5. So, the moral law cannot be externally imposed.
6. So, the moral law must be self-imposed; that is, autonomy requires that rational agents provide themselves with their own true moral principles. [*Self-Legislation*]

This argument shares its first premise and conclusion with premise (1) and (2), respectively, of *The Argument from Autonomy*. Along with the first premise, the second premise of *The Kantian Argument* also looks prima facie plausible. If externally-imposed and internally- or self-imposed are exhaustive categories, then this premise appears to express something like an analytic truth.¹⁰ However, once we get to the third premise, we can see that there is an ambiguity in the terms of the second premise that affects the way we read the argument.¹¹ What does it mean for the moral law to be *externally imposed*?

There appears to be an ambiguity here: *externally imposed* might either mean (a) that something external makes it the case that you φ or (b) that something external makes it the case that you *ought to* φ . Alternately, we should then suppose that *self-imposed* either means (c) that we make it the case that we φ or (d) that we make it the case that we *ought to* φ .

Now, if we substitute (a) into the third premise of the argument above, we get the following:

Yale—UConn Graduate Student Philosophy Conference (2009):
Kantian Constructivism and Self-Legislation

(3') If something external (i.e. the moral law) makes it the case that an agent φ -s, that rational agent could not count as acting on her own reasons.

This first substitution makes for a true premise, and on the whole the argument looks sound. For, if my action is determined by something outside of me and (let us also assume) I do not participate in my acting via any kind of deliberative process, then it seems plausible to say that I do not act on my own reasons. If the train's braking makes me lose my footing and fall onto a nearby passenger, I do not fall because I have decided to do so; rather, there is a sense in which the train's braking is responsible for my falling. Consequently, it seems safe to say that the reason for my losing my footing and falling is not my own.

It is also fairly uncontroversial that our concept of autonomy requires that we *do* act on our own reasons. That is, the fourth premise also appears true. If I were always just pushed around by physical and psychological forces without any kind of deliberative assent, I would not count as self-governing. So, if the moral law makes it the case that I in fact φ , then the existence of such a law would be incompatible with my own self-governance.

But, then on this interpretation, SL turns out to be an uninteresting conclusion. For if we substitute (a) into the third premise, we would then also need to substitute (c) into the conclusion. Here, the disambiguated conclusion merely states that the moral law alone does not force us to do anything; rather, we make it the case that we φ (where φ -ing, we will suppose, is some action-type prescribed by the moral law). But, *this* conclusion is not at odds with moral realism. For the moral realist need not claim that the moral law on its own is sufficient for getting an agent to φ . So, if this is the conclusion to

Yale—UConn Graduate Student Philosophy Conference (2009):
Kantian Constructivism and Self-Legislation

The Kantian Argument, the third premise of *The Argument from Autonomy* is false. It is not the case that if SL (so understood) is true, then moral realism is false. Moral realism is perfectly compatible with the claim that we make it the case that we φ , where φ -ing is some action type prescribed by the moral law. Although substituting the first reading makes *The Kantian Argument* sound, it also makes for an uninteresting conclusion. For, this conclusion does not do any work for the constructivist in the *The Argument from Autonomy*.

So, how do things look when we substitute (b) into the argument?

(3'') If something external makes it the case that an agent ought to φ , that rational agent could not count as acting on her own reasons.

The second substitution would make for an interesting conclusion. For if we substitute (b) into the third premise, we would then also be required to substitute (d) into the conclusion – viz., the reading of *self-imposed* that states that we make it the case that we ought to φ . This is a conclusion that would be at odds with moral realism. For the realist denies that the content of the moral law is *up to us* in a non-trivial way.

However, premise (3'') now looks to be false. For example, even if the rain makes it the case that Jones ought to bring an umbrella to the concert, this does not preclude Jones's acting on his own reasons when he decides to take his umbrella. Similarly, if Smith is to make her flight, she ought to arrive two hours early. Smith doesn't make it the case that she has this reason. But, nonetheless, it would appear that she acts on her own reasons if she decides to arrive two hours early, regardless of whether those reasons originate in her.¹² But, if this premise is false, then the argument turns out to be unsound – the appeal of its conclusion for constructivists notwithstanding. So, on the second reading, the conclusion may be interesting but the argument is unsound.

Yale—UConn Graduate Student Philosophy Conference (2009):
Kantian Constructivism and Self-Legislation

So, it looks as if *The Kantian Argument* fails to secure SL. It neither justifies the move from the first to the second premise in *The Argument from Autonomy* nor the status of SL as an independently-plausible thesis. So far then, there is no reason that we should accept the second premise of the constructivist's argument. The failure of this traditional argument for SL lends all the more importance to Rawls's alternative approach.

3. The Original Position and *Self-Legislation*

For those who are unhappy with *The Kantian Argument*, Rawls presents an alternative argument for SL. Considering that the traditional argument fails, the constructivist would appear to have a lot riding on the success of Rawls's argument. For, if the defender of constructivism is unable to marshal a viable argument in support of SL, she must forfeit *The Argument from Autonomy*. Rawls takes it that our intuitions about certain kinds of deliberative procedures support the interpretation of autonomy as *Self-Legislation*. Specifically, he claims that his *Original Position* [OP] is a procedure that bears this kind of connection with our concern for autonomy.

The OP is a hypothetical procedure by which citizens choose the principles that are to govern their society. A significant constraint on choice in the OP is the "the veil of ignorance", the condition that citizens must refrain from deliberating on information related to their particular identities and interests.¹³ Rawls claims that principles that result from such a choice procedure express the content of fair terms of cooperation and, hence, constitute a just social arrangement for a society.

In addition to determining principles, Rawls thinks that the OP helps us understand what autonomy requires and why it is important. He describes the OP

Yale—UConn Graduate Student Philosophy Conference (2009):
Kantian Constructivism and Self-Legislation

a *mediating model conception*¹⁴: it helps fix our ideas about autonomy by providing a framework that “...may enable us to crystallize our otherwise amorphous notion of the person and to identify with sufficient sharpness the appropriate characterization of free and equal moral personality.”¹⁵

The model conception of the OP is intended to balance two sets of judgments against each other. On the one hand, we have judgments about what kind of procedures fit with our considered judgments about fairness;¹⁶ on the other hand, we have judgments about “the appropriate characterization” of moral persons. Rawls is claiming that his description of the OP, with its veil of ignorance, strikes the right balance between these two sets of judgments. The OP is intuitively fair and part of what accounts for this fairness is the fact that the procedure is constrained in such a way that it allows rational agents the kind of freedom that constitutes their self-governance. But, what kind of freedom is this?

Rawls describes freedom as having different “aspects”. Here *aspects* can be understood as the more particular descriptions under which we recognize someone as free. One of these aspects is the description of a person as *a self-originating source of valid claims* (KCMT 543-4, 548).¹⁷

Freedom as a source of valid claims is a description of the kind of control over one’s life that is appropriate to a rational agent. On the one hand, it states that human beings confer value on their claims just in virtue of the fact that these claims are *theirs*. On the other hand, it states that the value conferred by such claims has implications for how rational agents are to act and that this is the only type of value that has this kind of implication for our actions.¹⁸

If our intuitions that the OP is fair are taken to support the idea that (i) rational agents are free to confer value on their own actions and (ii) this value is a source of claims, these same intuitions should support the idea that rational agents provide themselves with their own true moral principles – i.e., that they self-legislate. Insofar as we are the source of value, we are also the source of our own obligations.¹⁹ Rawls thinks that our intuitions about the fairness of the OP support the interpretation of freedom as a source of claims and, consequently, autonomy as SL.

4. The Self, Others, and Moral Constraint

There is a tension in the relation between the OP and SL, one that ultimately frustrates Rawls's attempt to make use of our intuitions about the OP in support of Kantian constructivism.²⁰ Rawls presents the OP as a choice procedure that is supposed to guarantee fair terms of social cooperation – terms that all rational and reasonable persons may accept. But, if the OP is to guarantee an outcome that will be fair to all such persons, the OP must constrain individual choice in such a way that this outcome is achieved. That is, the deliberators in the OP must have reason to constrain the pursuit of their own interests in order to facilitate the coordinated pursuit of goods and interests across a community of individuals.²¹ Let us call this the motivation of *Moral Constraint*.²² Although the description here is employed to determine principles of political justice, it is relevant for any moral theory. For any plausible moral theory must address the coordination problem and what justifying reasons an agent has to conform to other-regarding obligations.

However, if the OP merely allows for the freedom of a person as a self-originating source of claims, it does not look as if Rawls can account for *Moral Constraint*. Rawls wants to specify a property of autonomy whose full realization will result in an agent's being both self-governing (in the sense of SL) and moral.²³ But, it does not appear as if he can secure this result. On the one hand, if we assume that the full realization of autonomy makes an agent self-governing in the way that SL states, then the agent is only subject to those moral principles or laws which she has chosen herself. One may stipulate that agents will only have reason to choose principles that satisfy the demands of *Moral Constraint*. But, unless there is some further story to tell about why a self-governing agent would have reason to choose such principles, this move appears *ad hoc*. If there *is* some further story to tell, then it looks as if the principles that an agent chooses are constrained by externalities that she does not herself choose. For example, if the story is that all rational agents are sources of claims requiring respect from every other agent, then every agent is subject to at least some claims that she has not legislated for herself.

On the other hand, if we assume that the full realization of autonomy in Rawls's hypothetical procedure makes an agent moral, we do no better. First, this interpretation of autonomy stipulates an account of the reasons that would appear to require significant justification. One of the serious challenges for any moral theory is to explain why we have justifying reasons to perform other-regarding actions. To merely claim that our autonomy requires this is to put the cart before the horse. Second, if we start off with the demand to recognize and respect the claims that others make on us, then we are each

Yale—UConn Graduate Student Philosophy Conference (2009):
Kantian Constructivism and Self-Legislation

subject to a world of claims that we have not legislated for ourselves, not even hypothetically.²⁴

So, at best, it looks as if the OP does not support SL – at worst, as if the two are incompatible. The point here is not to rule out that the OP *models* any sense of autonomy; rather, the point is that if we accept that our considered judgments about autonomy do find expression in Rawls’s procedure, this autonomy cannot be SL.²⁵

5. Conclusion

I have canvassed two approaches that Kantian constructivists have employed to motivate the *Self-Legislation* thesis. Both of these approaches turn out to be self-defeating. *The Kantian Argument* is only able to secure the relevant thesis at the cost of a false premise and, ultimately, the argument’s soundness. Rawls’s argument from the OP presupposes a framework that proves at best unconnected and at worst incompatible with SL. The result of these failures is that the crucial premise in *The Argument from Autonomy* appears unsupported. Without this premise, Kantian constructivists must forfeit one of their strongest arguments against moral realism and the only one that uniquely picks out their view amongst competitors.

¹ John Rawls, “Kantian Constructivism in Moral Theory” (1980): 519

² See Gideon Rosen, “Objectivity and Modern Idealism: What is the Question?” in *Philosophy in Mind: The Place of Philosophy in the Study of Mind*. Eds. Michael, Michaelis and John O’Leary-Hawthorne. Boston, MA: Kluwer Academic Publishers, 1994. Pps. 277-319. As part of their case for the claim that their version of objectivity is objective enough, constructivists also argue that their view (1) allows that moral disagreement may be *genuine* disagreement and (2) helps us decide, in such cases, who is right. This point about disagreement is supposed to support the idea that the constructivist’s objectivity matches our considered judgments about objectivity. Again, I will not challenge the idea that the constructivist’s objectivity is objective *enough*. However, one might still reasonably object to the claim that the constructivist’s objectivity gets you *genuine* moral disagreement in *enough* cases. For, it has been charged that constructivism cannot block the prospect of an objectionable relativism (see Timmons 2003; also Johnston 1982 for a related point). This challenge may provide some real traction against Kantian constructivism. However, I will not be able to address this objection here.

³ This formulation is intended as a gloss. Further refinements introduce complications that would require more argument than I can afford here. Although this formulation is intended to be neutral between some interpretations of the view, it does already rule out others. One will note that, as I have stated the view here, Kantian constructivism is to be understood as a *metaethical* doctrine. However, it should be noted that it is disputed whether this interpretation best captures what ostensible defenders of the doctrine – viz., John Rawls, Christine Korsgaard, and T.M. Scanlon – have in mind. Some have argued that the doctrine is strictly to be understood as a first-order normative view that is quiet on metaethical issues. For recognition of metaethical constructivism in Rawls, see David Brink, *Moral Realism and the Foundations of Ethics* (New York: Cambridge, 1989), Appendix 4: 303-323; Ronald Milo (1995). For recognition of it in Scanlon, see Mark Timmons, “The Limits of Moral Constructivism,” *Ratio (new series)* XVI 4 December 2003, 391-423; Philip Pettit, “Review: Two Construals of Scanlon’s Contractualism” in *The Journal of Philosophy*, Vol. 97, No. 3 (Mar., 2000), pp. 148-164. For recognition of this view in Korsgaard, see William Fitzpatrick, “The Practical Turn in Ethical Theory: Korsgaard’s Constructivism, Realism, and the nature of Normativity” *Ethics* 115 (July 2005): 651-691. For recognition of quietism as a possible interpretation of Rawls, see Brink (1989); Stephen Darwall, Allan Gibbard, Peter Railton, “Toward a Fin de siècle Ethics: Some Trends” in *The Philosophical Review*, Vol. 101, No. 1, Philosophy in Review: Essays on Contemporary Philosophy (Jan., 1992), 115-189: 137-144. On this point concerning Scanlon, see Timmons (2003). Nadeem Hussain and Nishi Shah interpret Korsgaard as presenting her view under a metaethical guise, but they argue that this is not the best interpretation of her view; rather, they claim that Korsgaard is wrong to see her view other than as a version of quietism – see “Misunderstanding Metaethics: Korsgaard’s Rejection of Realism” in *Oxford Studies in Metaethics* 1, ed. Russ Shafer-Landau, (Oxford: Clarendon Press, 2006) 265-294.

⁴ Although defenders of Kantian constructivism present arguments that appeal to other moral concerns – e.g., concerns about the phenomenology of moral motivation and the putatively categorical authority of moral judgments, I take it that *The Argument from Autonomy* is the central argument presented in defense of Kantian constructivism. This is because it is the only argument which uniquely picks out the view. Arguments that start with a concern for moral motivation are compatible with expressivism, and those that appeal to a concern for the special authority of moral judgment appear to be just as well accommodated by other views that avoid all talk of autonomy and replace this with a more overt appeal to reasons (e.g. A particular act of φ -ing is *wrong* iff one has an all-things-considered reason not to φ in those circumstances). By contrast, Kantian constructivism is the only view that explicitly satisfies the demands presented by *Self-legislation*.

⁵ Here, it might be more apt to say that autonomy *makes the space* for morality possible.

⁶ Other relevant passages include G II 433; 440-445.

⁷ The constructivist line of interpretation is defended by Rawls (1980), “Themes in Kant’s Moral Philosophy” (1989), *History of Moral Philosophy* (2000); Korsgaard, *Sources of Normativity* (1996), *Creating the Kingdom of Ends* (1996/2000); Andrews Reath, “Legislating the Moral Law” (1994); J.B. Schneewind, “The Use of Autonomy in Ethical Theory” (1986), “Natural Law, Skepticism, and Methods of Ethics” (1991). This interpretation is rejected by Patrick Kain in “Self-Legislation in Kant’s Moral Philosophy” (2004); “Realism and Anti-Realism in Kant’s Second Critique” (2006); Karl Ameriks, *Kant and the Fate of Autonomy* (2000); *Interpreting Kant’s Critiques* (2003); Allen Wood, *Kant’s Ethical Thought* (1999); *Kantian Ethics* (2008); Terence Irwin, “Kantian Autonomy” (2004).

⁸ Such an appeal is clearest in Korsgaard’s discussion of autonomy: *SON* (1996): 97-98; *Creating the Kingdom of Ends* (1996, 2000): 24-27, 160-7; Rawls presents this interpretation of Kant in his *Lectures on the History of Moral Philosophy* (2000): pp. 226-230; 235-7; 252; 277-290. Much of this presentation is repeated from an earlier-published work, *Themes in Kant’s Moral Philosophy* (1989).

⁹ For example, it does not refer to the Kantian notion of the *will* as “a rational causality which is effective without being determined by any alien cause” (Korsgaard 1996: 97). It is also restricted to Kant’s argument for a negative conception of freedom; I omit the positive conception in the argument. Since the constructivist’s opponent here is the realist, the relevant notion of freedom concerns whether the moral law may be external to the rational will not whether there must be any law whatsoever.

¹⁰ One might object that this premise would have us presuppose that there is such a thing as the moral law and that, without further argument, it is not clear that we should accept this. In response to such worries, we might amend the premise, so that it looks like the following.

2'. If there is such a thing as the moral law, it is either self-imposed or externally-imposed.

This would also require us to us to modify premise (5) and our conclusion.

5'. So, if there is such a thing as the moral law, it cannot be externally imposed.

6'. So, if there is such a thing as the moral law, it must be self-imposed.

This conditional formulation of the conclusion would be weaker. However, one will note that if one takes 6' and plugs this in as the second premise of *The Argument from Autonomy*, the conclusion of that argument is not weakened in any significant way.

¹¹ My thanks to Matti Eklund for pointing out this ambiguity to me.

¹² One might object here that these examples describe prudential reasons – not moral reasons – and that this matters. For, in these cases, what Jones and Smith have reason to do will depend on what ends they have, ends that are in some sense *up to them*. If Jones has the end of making it to the concert in a condition that makes for comfortable attendance, he has a reason not to show up sopping wet. But, if he doesn't have this end – say, he enjoys sitting in wet trousers – he has no such reason. Similarly, Smith only has reason to show up at the airport two hours before her flight if it is important to her that she makes it to her flight on time and without stress. But, perhaps, she enjoys the thrill of rushing through security and running to her gate. In this case, she might not have any reason to show up two hours early. Importantly, whether Jones has the end of sitting high and dry during the concert or Smith has the end of relaxed travel is up to them, or, at least, it would appear so. The choice of such ends notwithstanding, it would still appear as if what Smith and Jones have instrumental reason to do is *not* up to them. This is what I hope to bring out with my examples. Let's say that Jones and Smith really do choose their ends, the former having the end of dry concert going and the latter the end of relaxed travel. Given these ends and facts about the world that constrain what means are available, Jones may not have a choice about whether or not he has reason to bring an umbrella; Smith may not have a choice about whether she has reason to arrive two hours early. Although each may exert some control in determining what reasons they have, the world pushes back. Jones and Smith determine the starting points for deliberation, but the world external to their choices makes it the case that Jones ought to bring an umbrella and that Smith ought to arrive two hours early for her flight.

¹³ They are asked, instead, to choose principles on the condition that they do not know who they would turn out to be in society – i.e. they do not know their social positions, or any of their particular inclinations, aspirations, or what their conceptions of the good would be (*ATJ* 16). Here, *conception of the good* means "...a determinate scheme of final ends and aims, and of desires that certain persons and associations, as objects of attachments and loyalties, should flourish" (*PL* 302).

¹⁴ 1980: 516, 520, 571-2. See also, Rawls's discussion of Kant's moral constructivism in *Lectures on the History of Moral Philosophy*, especially pp. 240-241 as well as this passage from *ATJ*, p. 226

¹⁵ 1980: 571-2

¹⁶ The description of the OP is supposed to line up with our considered judgments of fairness in the same way that descriptions of other procedures line up with judgments of fairness. For example, you-cut-I-choose is a fair procedure for allotting portions to which no one has a greater claim and we can recognize this independently of our theoretical commitments. The big difference, however, between this procedure for distribution and the OP is that there is an independent criterion we may appeal to for vetting our intuitions about fairness in the case of cake-cutting. In the case of you-cut-I-choose, the criterion is that no one should get more than any other person; the pieces should be as equal as possible. You-cut-I-choose guarantees this outcome (or, at least, an outcome that no one would have reason to complain about, given that no one deserves a larger share). But, the context of justice is importantly different in this respect. One cannot say in advance what a fair outcome would be prior to the agreement. Rawls wants to claim that it is the fact that principles would be chosen in the OP that makes the principles fair.

¹⁷ Rawls describes what it is to be a *self-originating source of claims* this way:

People are self-originating sources of claims in the sense that their claims carry weight on their own without being derived from prior duties or obligations owed to society or other persons, or, finally, as derived from, or assigned to, their particular social role. (KCMT 543)

It belongs to the parties' rational autonomy that there are no given antecedent principles external to their point of view to which they are bound. The interests they try to advance need not be derived from some prior duty or obligation, either to other persons or to society. Nor do the parties recognize certain intrinsic values known by rational intuition, for example, the perfectionist values of human excellence or of truth and beauty. This is how freedom as originating claims is represented. (KCMT 548)

¹⁸ Freedom as a source of claims is clearly plausible for explaining some instances of obligation. For example, we might reasonably assume that a piece of olive wood by itself does not have any special value; it is just a piece of olive wood like any other, with no more value than any other. But, if I take this piece of wood and make it into my wedding ring, it seems plausible to say that I give it a value that it did not have prior to the choice that gave it this place in my life. Surely, in the case of my ring, if someone asked to see it, took it from my finger, and then threw it off a bridge, I'd have reason to complain. This person would have wronged me; this action would be wrong because it does not express proper respect for the value I placed on my ring and the role that plays in my life.

¹⁹ This same point is echoed by Korsgaard in the third lecture of *Sources of Normativity*. For Korsgaard, human reflective consciousness creates a problem. Insofar as we can stand back and identify a self that is distinct from our impulses, we need a reason to act that is not yet another impulse. But where is this reason supposed to come from? For Korsgaard, the solution to this problem is that our conception of our practical identity – an identity that is ultimately founded on the value of our humanity – is the source of our reasons. That is, insofar as we are the source of our own value, we are the source of authority over our actions: “The reflective structure of human consciousness requires that you identify yourself with some law or principle which will govern your choices. It requires you to be a law to yourself. And that is the source of normativity” (1996: 104). See also Korsgaard (1996,2000): 239-43.

²⁰ The problem I draw out takes its inspiration from a dilemma presented by Jeremy Waldron in “Moral Autonomy and Personal Autonomy,” *Autonomy and the Challenges to Liberalism*, Eds. Christman, John and Joel Anderson. New York: Cambridge University Press, 2005: pp. 307-329.

²¹ Jeremy Waldron, “Moral Autonomy and Personal Autonomy.” pp. 307, 319

²² Rawls himself describes the agents in the OP as being *motivated* to accept certain types of agreements. He calls this aspect of their motivational set “reasonableness” or “a sense of justice”. I have opted to avoid talk of motivation. Instead, I restrict the discussion to a claim about what justifying reasons agents in the OP have. Rawls presumably thinks that there is a connection between reasons and motivation. But, I do not want to commit myself to an interpretation of what his precise views on this matter might be. As Thomas Baldwin has recently noted, Rawls's commitments in moral psychology are both complicated and understudied (“Rawls and Moral Psychology.” *Oxford Studies in Metaethics, Volume 3*. Ed., Russ Shafer-Landau. New York: Oxford University Press, 2008. Pp. 247-270).

²³ My thanks to Nick Sturgeon for helping me see this point in comments on an earlier version of this draft.

²⁴ What's more, if the legislation is merely hypothetical, then there does not appear to be any clear sense in which I have legislated the content of the moral law for myself. For, a hypothetical procedure may be one that I have never employed or even thought of.

²⁵ In fact, I argue elsewhere (MS) that the interpretation of autonomy that finds expression in Rawls's OP is metaethically neutral.