



Rawlsian Objectivity

ABSTRACT: *In a 1981 letter to H. L. A. Hart, John Rawls sketches a view of moral objectivity that substantially differs from that of contemporary constructivists. The view he describes does not rely on constitutive features of agency as Korsgaard's does, and it does not bottom out in a form of realism as Scanlon's moral theory does. Instead, Rawls's view grounds objectivity on the fundamental conceptions that could be shared in wide reflective equilibrium. Constructivism grounds objectivity in a kind of intersubjectivity, and Rawls finds the relevant kind of intersubjectivity in the alignment between fundamental convictions. This article develops this Rawlsian view of objectivity and highlights its strengths.*

KEYWORDS: John Rawls, moral objectivity, constructivism, relativism, constitutivism

In 1980, John Rawls sent a draft of his 'Kantian Constructivism in Moral Theory' lectures to his former teacher, H. L. A. Hart. In his response, Hart expressed general approval of the project but asked Rawls about his views on moral objectivity. Specifically, Hart asked whether Rawls believed that the meaning of 'you ought to do x' refers to objective reasons that one has (as Dworkin, Nagel, and Raz argued) or to more subjective facts (as Bernard Williams and Phillipa Foot argued). In reply, Rawls wrote, 'For the moment, I would interpret constructivism so as to be incompatible with Dworkin/Nagel/Raz and company', but he also wrote that he did not think that his view 'is what Foot and Williams hold either' (1981). Instead, Rawls gives a brief picture of a way in which moral judgments can have a kind of objectivity. Unlike that of contemporary theorists associated with Rawls, this kind of objectivity does not rely on conditions of agency or realism about reasons. Instead, it grounds objectivity on what fundamental conceptions could be shared in wide reflective equilibrium.

Following the publication of *Theory of Justice*, Rawls devoted much of his time to refining some of the deeper commitments he left unexplored in the book. According to his view, our moral principles are justified by a procedure that represents our nature as free and equal moral persons. In this argument, what ultimately justifies the principles of justice is the relationship between our 'fundamental conceptions' of moral persons, of the well-ordered society, and of the original position (Rawls 1999a: 303–58). Given this structure, Rawls determines the objectivity of the principles of justice by evaluating the objectivity of these supporting conceptions. Rawls asks Hart, 'are they objective in some sense?' If so, then the principles that follow from these conceptions can also be understood as objective.

Yet it seems odd to think that conceptions could be objective. After all, conceptions are not propositions, and it is not clear that their content could be

represented as propositions. In exploring the objectivity of these conceptions, Rawls writes,

Here I think there is no obvious path to a fruitful answer. It doesn't seem to help to ask the question in terms of the dichotomy between realism (moral truth) and relativism (or whatever). I believe that an illuminating answer must probe how many conceptions of the person there are and under what conditions of civilization and culture they flourish, or may exist, etc. If there is only one viable conception of the person that could flourish and be adopted as the basis of a political conception of justice in a society like ours (pluralistic, secular, large population, etc.) that provides a kind of basis for the conception of justice we affirm. (1981)

While this passage is quite opaque, it can guide us toward a distinctively Rawlsian view of moral objectivity. The task of this article is to develop that Rawlsian view.

What leads me to develop this account is ambivalence toward dominant forms of Kantian constructivism. While I am generally optimistic about the Kantian project, there is an open question about whether it places overly demanding conditions on agency. This concern gives me reason to look toward alternatives, which the Rawlsian view provides. Instead of relying on conditions of agency, the Rawlsian view grounds moral objectivity on a shared set of conceptions. My goal is to show how this view can provide a satisfying account of moral objectivity that is distinct from a constitutivist approach.

My arguments proceeds in four steps. In [section 1](#), I explain the constructivist approach to moral objectivity. I then explain how Rawlsian constructivism differs from the more familiar constructivist approaches in [section 2](#). In [section 3](#), I articulate the advantages of the view, and I address pressing objections in [section 4](#). Overall, the goal is to introduce a distinctively Rawlsian conception of objectivity that rivals both realism and constitutivism. In order to do so, I cannot here go into many of the details that a full metaethical account would require. I hope that the reader appreciates the breadth that precludes depth.

1. Constructivism and Objectivity

1.1

To begin, I will appeal to an intuitive distinction between the 'objective world' and the 'subjective world'. The *subjective world* is an individual's representation or experience of the world whereas the *objective world* consists in the world independent of such representations.

The world that exists independent of representation is often taken as a paradigm of 'objectivity'. Yet what is it that makes the objective world 'objective' in this sense? At first, it might seem like the objective world is objective because it is a world of *objects* (rather than a world of representations). Yet this is neither obvious nor helpful. First, it is not obvious because we would need to suppose that the

objective world itself consists of distinct objects rather than supposing that our cognition distinguishes parts of the world *into* objects. Second, it is not helpful because ‘consisting of objects’ does not get at what is significant about ‘objectivity’. When we appeal to ‘objectivity’ in conversation, our primary concern is not to refer to objects that are independent of representation. Rather, we typically mean to refer to some legitimate standard for determining the correctness of statements, propositions, or judgments. Our concern with the objective world is a concern with a world that serves as a standard of correctness and it is not primarily a concern with a world that consists of objects.

For this reason, I use the word *objectivity* to refer to that characteristic that a standard of correctness has. In common usage, words like *objectively* or *objective* are meant to refer to the absence of distorting prejudices or personal interpretations (e.g., ‘from an objective perspective’ or ‘look at the issue objectively’). I take it that what is left over after you have removed any distorting influences is whatever is correct.

For many judgments, the appropriate standard of correctness is the objective world. The correctness of the judgment that ‘it rained last night’ is determined by this standard, and the correctness of many judgments within the natural sciences are settled by the objective world. However, we do not need to think that the objective world is the appropriate standard of correctness for all judgments. For example, the judgment, ‘there is a largest prime number’ seems like it could be correct or incorrect, but it is not obviously a feature of the world of objects that would make it correct or incorrect. Likewise, the judgment ‘This water tastes like dust’ could, it seems, be correct, but it would depend on the experience of tasting rather than on what particles are in the water. Finally, the statement, ‘a statement is true insofar as it corresponds to the objective world’ could, it seems, be correct or incorrect, but it is not necessarily correct or incorrect because of the objective world.

There are other possible standards of correctness beyond the objective world, so objectivity is not fully determined by the world that exists independent of experience and representation. We should understand objectivity as more general than what is settled by the objective world. Objectivity is not necessarily a quality that the objective world uniquely has, so *moral objectivity* does not need to refer to something about the objective world that makes moral judgments true.

I.2

What is needed to make any conception of moral objectivity adequate is that it fulfills the role that objectivity plays in making sense of our daily activities. There are four activities, in particular, where a conception of objectivity plays a major role.

First, any act of judgment presupposes some standard of correctness. To form the judgment that it rained yesterday, one must presuppose that there is something that makes it correct or incorrect that it rained. Likewise, to judge that I acted wrongly in lying yesterday, one must suppose that there is something that makes it correct or incorrect that it was wrong to lie. By providing a ‘standard of correctness’, a

conception of objectivity thereby makes sense of the act of judgment because there is something that determines whether moral judgments are correct.

A second role that objectivity plays is to make sense of moral language. If you and I argue about whether it was wrong for me to have lied to your sister, we presuppose that there is a correct answer that we are arguing about. Moreover, Frege-Geach has taught us that we often embed moral propositions in more complex utterances, and making sense of these utterances (seemingly) presupposes that those propositions can be correct or incorrect. So, a conception of objectivity can make sense of our moral language by providing a standard of correctness for our disagreements and complex utterances.

Third, our practices of justification often presuppose some notion of objectivity. To see this point, we need to distinguish kinds of justification. The paradigm case of justification merely consists of person A giving a reason for their action to person B that B would accept. This paradigm does not presuppose objectivity because the reasons that someone offers are only intended to be accepted by B, and this does not require correctness. Outside of this paradigm case, however, we also offer justifications that an interlocutor *should* accept even when we do not think that they will. Here, we understand our interaction as appealing to objective standards of justification. Moreover, we also justify our choices to ourselves and write justifications that are not directed toward any specific person. An account of objectivity should be able to make sense of this second kind of justification.

Finally, there is also a kind of phenomenology that we would want an account of objectivity to explain. Just as we notice the difference between something being enjoyed and something being enjoyable, we might notice a difference between something being valued and something being valuable. We might also recognize a difference in feeling between (a) thinking that an action would be wrong and (b) having confidence that one is correct in thinking that the actions would be wrong. If this is right, then we would want an account of objectivity to make sense of this difference.

Of course, one might deny that there is any appropriate standard of correctness that underlies moral judgment, moral language, justification, or experience. The fact that we take ourselves to judge according to a standard of correctness does not mean that we actually do so. Imagine two people with blindfolds on, one that touches the back of a dog and the other that touches the back of a cat. The two might argue with one another about whether they have touched a cat or a dog and thereby take themselves to have touched the same animal. Yet this presupposition does not entail that they have actually done so. Likewise, we cannot conclude that there is any moral objectivity merely because our activities presuppose it. So, everything I have said is consistent with the denial of moral objectivity. The point of citing these activities is not to vindicate moral objectivity, but better to articulate what we would want some account of moral objectivity to do.

Someone who offers a conception of moral objectivity, as I do here, offers an account of the standard of correctness that underlies moral judgment, language, justification, and experience. So, we should not assume that a conception of objectivity must account for mind-independent moral properties. To claim that would beg the question in favor of one conception of moral objectivity over

others. Instead, different conceptions of objectivity are rivals insofar as they make sense of these four activities in different ways.

I.3

The constructivist recognizes that moral judgments can be correct or incorrect, but their correctness is not determined by the objective world. In this way, the constructivist does not need to claim that the objective world has any moral properties. Instead, whether a moral statement is correct is understood as a relation between the statement and the outcome of a constructed procedure or—to use Sharon Street’s formulation (2010)—the relation between a statement and a practical perspective. If we interpret Kant as a constructivist, for example, the statement ‘you morally ought to do X’ is correct if a maxim that avoided doing X could not be willed as universal law. According to Scanlon’s constructivism, the statement ‘you owe others your doing X’ is correct when the action accords with a principle that no one could reasonably reject. According to Rawls’s constructivism, ‘the basic structure of society ought to meet condition X’ is correct when persons would agree to condition X in the original position.

With this conception of objectivity, the constructivist differs from salient kinds of moral realism and subjectivism. First, if we define *moral realism* as the view that the objective world determines the correctness of moral judgments, then the constructivist is not a moral realist. The outcome of the procedure is what determines correctness and not features of the objective world. Second, if we define subjectivism as the view that there is no valid standard by which to judge the correctness of moral judgments, then constructivists are not subjectivists. They recognize a standard for correctness in the outcome of their procedure.

I.4

This leaves the constructivist with a particularly difficult challenge: in virtue of what can the constructivist identify the outcome of their preferred procedure as the appropriate standard for correctness? They cannot say that a feature of the objective world shows that the procedure is appropriate, and they cannot say that the procedure itself warrants the procedure. So what licenses the constructive procedure as determining the standard of correctness? I take this to be the most difficult question that the constructivist must answer.

Different constructivists will argue for their preferred procedures in different ways, but there is one similarity between these approaches. Instead of arguing from some feature of the objective world, they argue that their preferred procedure matches up to the subjective world *for each and every person*. An appeal to the ‘practical perspective’, for example, evokes a feature of subjective experience for each and every agent. In this way, constructivists identify some feature of individual experience or representation that is related to what I call ‘the problem of ethics’. Then, they argue that this feature of individual experience serves as an adequate basis for correctness only because it is a feature of experience or representation for each and every person. Accordingly, the constructivist argues

for their preferred procedure by showing that (a) it uniquely relates to the problem of ethics, and (b) relates to the problem for each and every person.

In a slogan, the constructivist takes intersubjectivity as the basis for objectivity. The objective world is objective in the sense that it serves as the appropriate standard for judgments about the world independent of experience and representation. The outcome of a constructive procedure is objective in the sense that it serves as the appropriate standard for moral judgments. The constructive procedure is linked to the problem of ethics for each and every person, so it is intersubjective. Just as Miriam Ronzoni emphasizes the link between constructivism and intersubjectivity in ‘Constructivism and Practical Reason’ (2010), I argue that objectivity is grounded by intersubjectivity for the constructivist.

2. Kantian and Rawlsian Constructivism

2.1

Of course, this constructivist project fails if it cannot show why a particular procedure has intersubjective validity. The constructivist needs to show (a) that her procedure connects to ‘the problem of ethics’ and (b) why it connects to the problem of ethics for each and every person. Yet, how could this intersubjective standard be possible? The Rawlsian view gives a very different answer to this question than contemporary Kantian constructivists do.

Kantian accounts claim that their standard of correctness is related to the problem of ethics because of the conditions of agency. For example, Korsgaard aims to show why valuing humanity is constitutive of agency, so each and every person—as an agent—values humanity. For this reason, we can identify the standard of correctness for moral claims as correspondence with valuing humanity (2009). By identifying what is constitutive of agency, the Kantian connects a standard of correctness with the problems of ethics for each and every person.

By contrast, Rawlsian constructivism does not rely on claims about agency. Instead, Rawls’s constructive procedure grounds the standard of correctness on an agreement between fundamental conceptions. These conceptions are not necessarily linked with any comprehensive doctrine or philosophical theory but represent our basic understanding and core convictions about the world, our society, our relationships, and ourselves. If we share a conception of the person or society, for example, then the constructive procedure generated from those conceptions has intersubjectivity. By identifying the standard of correctness with the results of this procedure, the Rawlsian achieves a kind of objectivity as intersubjectivity. The constructive procedure is related to the problem of ethics because ethics seeks to bring our actions into accord with our convictions, and our fundamental conceptions express our convictions. When persons share fundamental conceptions of the person and society, then the judgments that are entailed by those conceptions have a kind of intersubjective validity.

In short, Rawlsian objectivity is an intersubjectivity of fundamental conceptions. The whole structure of the Rawlsian argument unfolds from a commitment to these

conceptions, and the results of this argument are intersubjectively valid insofar as each and every person can accept these conceptions.

2.2

At first glance, this Rawlsian view seems to be wrong. Most persons do not have any articulable conception of the person or of society, and those who do have such conceptions rarely agree with one another. Accordingly, it seems there are no conceptions of the person or society that could be intersubjectively valid. So, Rawls would be unable to get objectivity.

However, such an argument ignores an important nuance in the Rawlsian view. For there to be a standard of correctness, there must be some accord between fundamental conceptions. Yet, there doesn't need to be an accord between current conceptions. Instead, there only needs to be an accord between those conceptions that persons *would* hold in wide reflective equilibrium. The idea is that if persons scrutinized and revised their judgments, they would come to accept that these conceptions of the person and society fit with their deeper convictions. For example, the principles of justice would have a kind of objectivity if persons would accept the Rawlsian conceptions of the person and society in reflective equilibrium.

If this is correct, it represents a major switch in understanding what grounds objectivity in Rawlsian theory. It is wide reflective equilibrium—rather than constructivism on its own—that provides the ultimate basis for objectivity as intersubjectivity. This does not change the fact that the principles of justice are valid only if they are warranted by Rawls's constructed procedure, but the procedure only licenses principles of justice because the conceptions—from which the procedure is constructed—have objectivity.

I am not the first to make the observation that reflective equilibrium provides the foundational justification for our convictions in Rawlsian theory. In 'Constructivism as Rhetoric', Anthony Simon Laden provides an interpretive argument that follows Rawls's remarks from his earliest to his latest writing (2013). He concludes that the constructive procedure is best understood as a rhetorical device that shows the relationship between core convictions in a way that motivates the reader to accept Rawlsian principles. Regardless of whether Rawls's constructivism is primarily a rhetorical device, Laden's argument makes it clear that reflective equilibrium is more foundational for justification than any constructive procedure. After all, the constructive procedure should be significant to the reader only if it accords with her conceptions of the person and of society upon due reflection (Rawls 1999b: 18).

Rawls's letter to Hart provided additional support to Laden's view. There, Rawls claims that the objectivity of the principles of justice depends on the objectivity of conceptions. He cannot have meant that these conceptions are objective if they are warranted by a procedure because the procedure is constructed from these conceptions. Instead, the conceptions are 'objective' if they are the conceptions that persons would hold in wide reflective equilibrium.

2.3

This Rawlsian view puts fewer demands on agency than the Kantian view, but what would justify us in thinking that persons would come to an interpersonal accord in their conceptions? It can seem merely hopeful to suppose that our conceptions of the person and society would align in reflective equilibrium.

To address this issue, we need to distinguish two questions: (1) do our conceptions align, and (2) is it possible for our conceptions to align. We can put the first question aside because it can only be answered by going through the long process of achieving wide reflective equilibrium. In this way, the questions can only be answered through the arguments and considerations that moral theory involves. We cannot know what the end of this process would be without going through the process itself, so we cannot expect to answer the first question now. For an account of objectivity, we do not need to know what is correct or incorrect, we only need the existence of a standard of correctness. That standard is what is entailed by the fundamental conceptions people would agree on in wide reflective equilibrium, regardless of whether we can currently identify what that is. However, this project would be imperiled if an accord between our fundamental conceptions was shown to be impossible. So, we need to address the second question. What grounds do we have to think such an accord is possible?

In his letter to Hart, Rawls wrote, 'if there is only one viable conception of the person that could flourish and be adopted as the basis of a political conception of justice in a society like ours (pluralistic, secular, large population, etc.) that provides a kind of basis for the conception of justice we affirm'. On first look, this seems like a very strange remark. Why would the conceptions that are likely to ground a conception of justice in contemporary society provide a basis for justice? The remark is not strange, however, if it is intended to answer, 'is it possible for our conceptions to align?' There could be an accord between fundamental conceptions because there are only a few fundamental conceptions that could provide a kind of basis for justice in a society like ours.

In this remark, Rawls appeals to the resources that are familiar from a long line of theorists from Montesquieu, Rousseau, Hegel, Marx, and Wittgenstein. Our social atmosphere influences our understanding, values, and desires. If the features of our social atmosphere are of a particular kind, they can lead those who live in that social atmosphere toward certain conceptions of the person and society. For both Rawls and Wittgenstein, our ways of thinking are often socially determined by the form of life in which we live. The fact that we live within a particular social atmosphere ('pluralistic, secular, larger population, etc.') is likely to lead to particular ways of thinking about ourselves and the world around us. Rawls offers a similar remark in the published version of 'Kantian Constructivism in Moral Theory' (1999a: 355). Insofar as we share a way of life, it is conceivable that we will share substantive conceptions that are influenced by that social atmosphere. While it remains only hopeful that these conceptions can serve as the basis for interpersonal accord, we have reason to think why it is possible. The conceptions can be intersubjectively valid because they are influenced by a shared social world.

2.4

So far, I have contrasted the Rawlsian view of objectivity with Kantian accounts like Korsgaard's. Now, I will contrast the Rawlsian view with Sharon Street's Humean view. Most starkly, the Rawlsian view identifies moral objectivity with intersubjectivity whereas Street identifies objectivity with subjectivity. For Street, someone has a reason to do something when such a reason is entailed by their set of normative judgments. For the Rawlsian view, someone has a reason to do something when such a reason is entailed by the fundamental conceptions that people would share in reflective equilibrium. This does not mean that every pursuit would need to be deemed worthwhile in wide reflection for people to have reason to pursue it. Rather, it is quite likely that our shared fundamental conceptions would entail that people have reason to pursue what they want, value, or desire. Nonetheless, the Rawlsian view is more restrictive in what people have reason to do than Street's view would be.

For both the Humean and Rawlsian view, the standard of correctness depends on the contingent set of convictions that agents hold. The difference is that the Humean account does not require that the standard of correctness for normative judgments be consistent across different people. For each individual, it is their own set of convictions that determines whether they have reason to do something. For the Rawlsian view, what ultimately matters is the set of convictions that people share. In this way, the Rawlsian view provides a clear contrast between (a) valuing something and (b) judging it to be valuable. For the Rawlsian, to value X is to have some set of attitudes and dispositions toward X. By contrast, to make the judgment that X is valuable is to believe that—according to the standard of correctness—it is valuable. This judgment would be correct if the convictions people would share in reflective equilibrium would entail that it is valuable. By contrast, the distinction between valuing something and judging it to be valuable is less distinct in the Humean account because whether something is valuable is determined simply by whether one values it.

The Rawlsian view developed here tries to balance between more Kantian and more Humean versions of constructivism. While the Humean account is concerned with the convictions that persons would have under ideal conditions, it does not provide a standard of correctness that is shared across persons. While the Kantian account provides a standard of correctness that is shared across persons, it does so on the basis of contentious features of agency. As I discuss in section 3.3, what follows from these conditions of agency might not align with a person's convictions, and that may open a rift between what we are committed to and what we objectively should do. Alternatively, the Rawlsian view aims to build a conception of objectivity grounded on the convictions that agents in reflective equilibrium would share. This gets a shared standard of correctness that is necessarily tied to each agent's own commitments.

2.5

If objectivity is understood as a form of intersubjectivity, then we need to identify the group among whom intersubjectivity is required. For a standard of correctness,

which group of people would need to share fundamental conceptions in reflective equilibrium?

Under one view, this group should include all beings that are capable of understanding reasons for action. So, a second-century Roman and twenty-second-century Klingon would need to come to intersubjective accord for objectivity. While there is nothing that makes such an accord impossible in principle, we could not rely on shared circumstances or a way of life to ground intersubjectivity. However, it might still be possible to identify some feature of agency as the basis of intersubjectivity between diverse people. This approach would still differ from typical versions of Kantian constructivism because it would not need to ground moral objectivity on the *necessary* features of agency. Instead, it would require that people happen to share some conceptions because of the experience of being an agent. Instead of grounding intersubjectivity in the *constitutive* features of agency, intersubjective agreement would depend on *contingent* features of agency. Whether there are these contingent features would be hard to prove, but it is nonetheless a possible ground for an intersubjective agreement across all beings capable of engaging in practical reason.

A second possible answer would identify the relevant group as all those persons whose lives we can affect through our choices. Here, the idea is that we owe justification for our actions to those we affect, so we should seek some accord between our conceptions and theirs. Yet this would require accord with persons in the distant future whose way of life might be unimaginably different. Since second-century Romans could affect the lives of twenty-second-century Klingons, this second answer would not be substantively different from the first.

A third, and more promising, response identifies a less extensive group. It supposes that objectivity requires intersubjective agreement between those who can exchange reasons. Recall the role that objectivity plays in moral judgment, moral language, justification, and experience. If we want an account of objectivity to vindicate what we take ourselves to do when we judge, converse, and justify, then we need that account of objectivity to be interpersonally valid between those that we judge, converse, and justify *with*. This condition does not require that intersubjective agreement is between all reasoners, but merely between those that engage with one another. To provide an adequate account of objectivity—wherein adequacy is measured by the ability to vindicate the activities that presuppose some notion of objectivity—the group of individuals that intersubjective agreement is needed between is the group of people who converse and justify with each other. For short, I say that there needs be intersubjective agreement within a ‘justificatory community’.

With this third understanding of the relevant group, the experience of being an agent is not the only resource that explains why intersubjective agreement might be possible. Instead, we can appeal to the contingent features of a shared way of life. We do not need to be concerned with what fundamental conceptions all people would share because all people do not engage in moral conversation and justification with each other. However, our community does engage with most of the world at this point. Even if I live my life in an insular town, those who I engage with engage with others who engage with those around the world. So, the

group among whom we seek intersubjective agreement is currently quite broad and diverse. Nonetheless, there seems a greater chance that this group shares some features of a way of life that we would not share with all agents.

Given his remarks, it seems the third understanding is the one that Rawls would have in mind. His focus on ‘what conditions of civilization and culture’ a conception of the person could flourish in, suggests that the fundamental conceptions that ground objectivity were specific to a limited time period. Yet, much of the theory expressed here is consistent with the first answer. If features of agency—either necessary or contingent—explain why persons would share fundamental conceptions, then that provides a basis for agreement in fundamental conceptions that goes beyond the conditions of civilization and culture.

2.6

The Rawlsian view starts by recognizing the nature of moral argument as a disagreement between convictions. In seeking moral objectivity, we seek some standard of correctness to settle this disagreement. The Rawlsian conception of objectivity identifies that standard of correctness by appeal to what conceptions people would share in wide reflective equilibrium. From that point of agreement, we can build further and more substantive agreement. For example, from agreement about a conception of the person and of society, we can construct a procedure that respects our points of agreement. The outcome of that constructed procedure would then be valid principles of justice. In this way, Rawls identifies the point from which such justifications could proceed as fundamental conceptions. Moral principles are objective insofar as they follow from commitments to these conceptions. To put the Rawlsian view into a slogan, ‘moral objectivity consists in a justification that all persons could accept in reflective equilibrium’.

3. Advantages of the Rawlsian View

3.1

While this picture might not provide what some readers expect from an account of objectivity, it does offer what we would need a conception of moral objectivity to do. In section 1.2, I argued that moral objectivity has a role in making sense of moral judgment, moral language, practices of justification, and phenomenology. The Rawlsian view provides a conception of objectivity that fulfills these roles.

First, it can make sense of moral judgment because there is a standard of correctness that we can judge claims to correspond with. While it might be a long process to go from shared fundamental conceptions to such particular judgments, this account provides a basis for that judgment. In judging that I should not lie to your sister, I am supposing that it is either the case or not the case that lying to your sister is inconsistent with shared fundamental conceptions.

For similar reasons, the account can also make sense of moral language. We can embed moral claims within complex utterances because they can be correct or

incorrect. In saying, 'If I should not lie to your sister, then I should not lie to my brother', I am not committed to either judgment about what I should do but am only expressing a relationship between the claims. This account provides a standard of correctness that can make sense of this relation.

Likewise, if I either justify my decisions to you on the basis of reasons that you do not (but should) accept or justify choices to myself, I can make sense of reasons that I judge to be correct. I am not merely trying to appeal to what you or I care about but I appeal to what is objectively important.

Finally, this account can make sense of the phenomenology of objectivity. For example, we can explain the difference between valuing something and believing it to be valuable by appeal to the belief that it is morally correct to value it. The belief makes sense as a belief about what fundamental conceptions people would share in reflective equilibrium, or a belief about what would be entailed by those conceptions.

At this stage, I am not sure what else to say to someone who expected something different from an account of objectivity. If this view can accomplish everything required of objectivity, then it seems to be exactly what one should expect. It does not provide a mind-independent standard of correctness, but there is no question-begging reason to insist that it should. The account does not ground morality in constitutive features of practical reason, but there is no non-question-begging reason to insist that it should. What matters is that the account explains the role that objectivity plays in our understanding and activities.

3.2

Beyond satisfying the requirements of an account of objectivity, one major advantage of the Rawlsian view is that the correct moral judgments are always those that we would accept on due reflection. So, there is no way to identify the correct moral principles independent of ethical reflection. Thus, the Rawlsian view of objectivity supports the independence of moral theory.

If I maintain that Rawls's two principles of justice are correct moral principles and I accept this conception of objectivity, then I believe that persons would agree with the Rawlsian conceptions of the person and society in wide reflective equilibrium. If you disagree, then I can only show that I am correct by convincing you to accept these conceptions. After all, whether the principles are correct depends, in part, on whether you would accept them. Likewise, you should try to convince me that these conceptions are not acceptable (or that specific judgments do not follow from them). Our argument is always about what our fundamental commitments should be. The truths of physics, psychology, and biology are irrelevant unless they affect our convictions. So, the only way to identify a standard of correctness is by engaging in moral theory.

In this way, the conception of objectivity supports—what Rawls called—the independence of moral theory. Rawls argued that moral theory should be independent of epistemology, philosophy of language, or philosophy of mind. In making this argument, Rawls avoids the eighteenth-century tendency to build all of philosophy from a theory of knowledge or the twentieth-century tendency to

build from a theory of meaning. Instead, ‘moral theory is the study of substantive moral conceptions’ that is studied by trying ‘to find a scheme of principles that match people’s considered judgments and general convictions in reflective equilibrium’ (1999a: 286, 288). Other parts of philosophy or science do not have explanatory priority because these other areas do not dictate what people would accept in reflective equilibrium.

The Rawlsian view of objectivity better explains why it would not be worthwhile to ground a moral theory on a theory of knowledge, language, or mind. First, the account does not suppose that there is some set of moral facts, which we must find reliable access to. The only way to identify the standard of correctness is ‘to find a scheme of principles that matches people’s considered judgments and general convictions in reflective equilibrium’ (1999a: 287). In other words, the only way to determine how we ought to act is to engage in the ‘study of substantive moral conceptions’. So, moral theory can proceed independently of epistemology. Likewise, the object of discussion when using moral language is the set of judgments that are entailed by the fundamental conceptions we would share in reflective equilibrium. While philosophy of language is needed to make sense of this discussion, it does not settle what conceptions we would share. So, moral theory proceeds independently of a theory of meaning.

For similar reasons, moral theory is independent of the truths of physics, biology, and psychology. The truths of these fields will likely bear on what persons could accept, but the information is only relevant insofar as it impacts reflective equilibrium. In this way, the Rawlsian view of objectivity vindicates a common intuition: we should go about determining what is morally correct by engaging in moral theory.

3.3

Since the Rawlsian view supports the independence of moral theory generally, it also shows why moral principles do not depend on a philosophy of action. As such, the Rawlsian view avoids David Enoch’s recent criticism of constitutivism (Enoch 2006).

Constitutivism grounds certain reasons for action on the norms that constitute agency. If valuing humanity is necessary for one to be an agent, then persons have reason to value humanity. If self-interpretation is a necessary feature of agency, then persons have reason to act in ways that make sense to themselves. Enoch’s criticism of constitutivism does not dispute that there are norms such as these that constitute agency. Instead, it doubts that such norms provide any reason to act. To see the point, imagine a fully convincing argument that shows how the requirements of agency entail that persons should do X. Thus, anyone who fails to do X would not be acting. Enoch grants this, but challenges its significance. If someone is otherwise against doing X—or just does not care about whether they do it—it is open to them to deny being an agent (Enoch 2006). One might say, ‘Okay, you have convinced me that I am not an agent because I fail to do X. That is fine. Instead, you can call me a “schmagent”. If I am not engaged in agency, I am engaged in schmagency. It is like agency in many ways, but it is not agency because it does not involve doing X’.

This ‘schmagency objection’ denies the significance of constitutivism by opening a rift between the conditions of agency and one’s commitments. If I am committed against doing X and I see that agency requires me to do X, then I can give up on any commitment to agency. I will live my life in ways that are similar to being an agent except that I will not do X. The constitutivist account relies on facts about agency to provide authoritative reasons for acting. Enoch’s objection does not challenge whether these facts about agency are true. It challenges the moral authority of such facts.

Regardless of whether this criticism succeeds against constitutivism—and there is much debate on that question—it does not succeed against the Rawlsian view. The Rawlsian view does not rely on constitutive norms of agency to ground reasons to act. What makes the standard of correctness authoritative is that persons would accept it in reflective equilibrium. So, there is no possibility of a rift between one’s committed convictions and what one objectively should do. If anyone were to give up on how they objectively should act, it is only because they fail to recognize what they themselves would be committed to do if they reflected more. This is not because everyone is necessarily committed to some substantive view, but because what makes the objective standard objective is that everyone would be committed to it.

3.4

As a final advantage, the Rawlsian view is also fully ‘naturalist’. Whether or not there is a clear and satisfying articulation of what naturalism entails, there are certainly intuitive standards. Naturalists avoid appealing to any ‘spooky’ properties that are not open to observation and testing. A commitment to naturalism, for example, eschews causal explanations that appeal to ghosts, gods, and souls. Often, naturalism is also taken to eschew causal explanations that appeal to mind-independent moral properties. Conceived in this way, Rawlsian objectivity is fully consistent with naturalism. What grounds the fundamental standard of correctness is our convictions, which are a familiar part of everyday life. The notion of agreement between fundamental conceptions does not require any spooky properties. It does not require ghosts, gods, souls, or mind-independent moral properties. So a committed naturalist is free to accept the Rawlsian view of moral objectivity.

4. Objections to the View

4.1

Since the account relies on what persons would accept in reflective equilibrium, we might worry that another conception of objectivity is sneaking in through the process of reflective equilibrium. Is not the process toward reflective equilibrium a norms-guided process? What determines the norms that specify that process? We might worry that we are only getting intersubjective agreement in reflective equilibrium because of how we stipulate what makes for reflective equilibrium.

There are two versions of this criticism to distinguish. The first version sees reflective equilibrium as a good goal for reasoning because it reliably gets the correct beliefs and convictions. According to this view, the attempt to bring our beliefs and convictions into alignment is worthwhile because doing so tends to get us the truth. If there is any dispute about how the process of reflective equilibrium should go, that dispute would be settled by which process is more likely to reliably issue in correct beliefs and convictions. If we conceive of reflective equilibrium in this way, however, then the Rawlsian account would be self-defeating. What justifies using reflective equilibrium would be its reliability in getting us the correct beliefs, so what makes a belief correct would need to be determined independently from the process of reflective equilibrium.

This first version of the criticism misunderstands the role that reflective equilibrium plays in the Rawlsian view of moral objectivity. It is not intended to be a reliable process to identify what is independently true. Instead, wide reflective equilibrium determines the correct moral beliefs and convictions (see Scanlon 2003; Freeman 2007: 29–42, Daniels 1980). In the Rawlsian view, there is no standard of correctness beyond what convictions people in reflective equilibrium would share. The process of reflective equilibrium cannot be justified as identifying those independently correct beliefs and convictions. While this conception of reflective equilibrium is clearly incompatible with epistemic justification, Samuel Freeman emphasizes that ‘Rawls sees reflective equilibrium as particularly a claim about moral justification, not justification in general’ (2007: 41). So, someone who presses the first version of the criticism is merely confusing the role that reflective equilibrium plays in the theory.

The second version of this criticism claims that reflective equilibrium requires norms that regulate the process even when it constitutes a standard of correctness. We assume that there are such norms if we suppose that someone can go through the process of reflective equilibrium incorrectly. If they can, then there must be a set of norms that guide the process. Yet, what would make them the *correct* norms to regulate the process of reflective equilibrium? To identify the correct norms that guide the process, we would need some standard of objectivity that is independent of the results of reflective equilibrium. Thus, recognizing norms for reflective equilibrium would require that some other standard of objectivity be smuggled in.

Yet, this second criticism goes wrong in supposing that there must be some correct set of norms that regulate reflective equilibrium. What matters about reflective equilibrium is that a person comes to a point where they can endorse their set of beliefs and convictions as a whole. We do not need a set of norms to regulate how people get to that point. Once someone is there, all the various aspects and entailments have been considered and endorsed. What matters is their being at this point and not how they got there. While we often associate a process of reflective equilibrium with bringing our convictions into coherence, we do not do so because it essentially demands coherence. Rather, we do so because we would not endorse a set of beliefs and convictions that do not cohere. What is doing the work is the aim of achieving reflective endorsement, and not whether a set of guiding norms is correct. So we should resist any set of norms that regulates the

process of achieving equilibrium. A person might fail to achieve equilibrium, but they would not be going through the process wrongly. For this reason, the second version of the criticism is unproblematic as well.

4.2

In 2.5, I argued that there might be different justificatory communities that agree on different fundamental conceptions. Since this agreement between conceptions establishes a standard of correctness, I have opened the conceptual possibility of a kind of moral relativism. There could be a justificatory community that shares an agreement in fundamental conceptions, and there could be a disconnected second community that shares a different agreement in fundamental conceptions. In this case, the standard of correctness for one community would be different from the standard of correctness for the other. What is correct would then be relative to the justificatory community that one is situated in.

Yet I do not see how this conceptual possibility is a problem for an account of objectivity. First, the possibility does not undermine the role that objectivity plays in our activities. Any particular individual would still have a standard of correctness for their judgments, and that standard of correctness would be interpersonally valid among those that use moral language and justify their actions with it. It is only if we presume that objectivity is inconsistent with the possibility of relativism that this would be a problem, and I do not see a reason to presume that.

Second, the conceptual possibility of relativism does not even mean that relativism holds. We might suspect that the set of possible intersubjective accords is quite limited by features of the human experience. It might be unlikely, for example, that any justificatory community could ever agree on the superiority of a particular sex.

4.3

A third objection is motivated by a broader interpretation of Rawls's work. In *Political Liberalism*, Rawls distinguishes 'true moral principles' from 'a political conception of justice'. While true moral principles are based on comprehensive doctrines, a political conception of justice depends on an overlapping consensus. In other words, a political conception of justice is grounded on reasons that all persons—with their different comprehensive doctrines—can accept for regulating the basic structure of society. In this way, both the standard of justice in *Political Liberalism* and the standard of correctness developed here rely on shared convictions. A political conception of justice is appropriate because citizens share reasons and a standard of moral correctness is appropriate because people share fundamental conceptions.

Given this similarity, one might be tempted to think that the view of objectivity that Rawls describes in his letter to Hart was only intended to ground a political conception of justice. The third objection to this account claims that my reading of Rawls goes too far in supposing that he offers a general conception of moral

objectivity instead of a political conception of justice. This objection is made even more plausible once we recognized that Rawls's letter to Hart was written in 1980, so the remarks could represent an earlier stage in the development of his ideas. Someone might suppose that Rawls never saw shared conceptions as an adequate ground for objectivity *tout court* but did see them as a fully adequate ground for a standard of political justice.

In reply to this objection, I first give a conciliatory response and then a combative one. My conciliatory response supposes that the objection is correct. Suppose that Rawls really only intended to discuss a political conception of justice in his letter to Hart, and my remarks go too far in applying them to a full account of moral objectivity. That would mean that I am not articulating Rawls's view of moral objectivity. Nonetheless, I am articulating a Rawlsian view of moral objectivity. My primary goal is to articulate a plausible view that is aligned with Rawlsian thought, and that goal can be accomplished regardless of whether it was or was not Rawls's own final view.

For a more combative response, I offer evidence that this account of objectivity is actually what Rawls had in mind. After all, *Political Liberalism* is not the only place where Rawls discusses objectivity. In his 'Lectures on the History of Moral Philosophy', Rawls describes the Kantian conception of objectivity as follows:

Moral convictions are objective if reasonable and rational persons who are sufficiently intelligent and conscientious in exercising their powers of practical reason would eventually endorse those convictions, when all concerned know the relevant facts and sufficiently surveyed the relevant considerations. To say that a moral conviction is objective, then, is to say that there are sufficient reasons to convince all reasonable persons that it is valid or correct. (2000: 245)

Kant was not merely articulating a political conception of justice or, even, a reasonable basis for social morality. Kant offers a conception of moral objectivity, yet Rawls describes that conception as relying on intersubjectivity in the same way as the Rawlsian view defended in this article. In offering this Kantian view of objectivity, Rawls focuses on the fact that all persons *would endorse* the basic Kantian conceptions. This is also true of the Rawlsian conception of objectivity. When Rawls contrasts Kantian Constructivism with rational intuitionism in 'Kantian Constructivism in Moral Theory,' (1999a: 355)' and 'Themes in Kant's Moral Philosophy' (1999a: 515–16), he specifically discusses this kind of agreement as a ground for objectivity. The key difference between Kant's view of objectivity and Rawls's view is that Rawls did not rely on necessary features of agency to explain why persons would endorse these conceptions. Instead, contingent features of experience explain the possible alignment in conceptions, whether those contingent features are typical of agents or typical of those in our justificatory community.

Rawls also suggests such a view of objectivity in his 1975 article, 'The Independence of Moral Theory'. When reviewing his argument, he writes,

I first discussed the method of reflective equilibrium and suggested that the questions as to the existence of objective moral truths seems to depend on the kind and extent of agreement that would obtain among rational persons who have achieved, or sufficiently approached, wide reflective equilibrium. (1999a, 301)

Here again, Rawls suggests that moral objectivity—not a political conception of justice—depends on the extent of agreement in wide reflective equilibrium. He did not commit himself to such a view of objectivity in this 1975 article, but he was enthusiastic about its plausibility. Even in his few remarks about objectivity in *Political Liberalism*, Rawls regularly cites an essay by Warren Quinn that offers a general account of objectivity that appeals to a device like wide reflective equilibrium (Quinn 1987: 199). Since Rawls did not discuss his views of moral objectivity frequently, and passages like these suggest a view like that discussed here, I do not think we can dismiss his remarks as merely a precursor to the views of *Political Liberalism*.

Moreover, the aim of an account of objectivity is substantially different from the aim of *Political Liberalism*. In ‘Kantian Constructivism in Moral Theory’ and the letter to Hart, Rawls aims to show that rational intuitionism is not necessary for moral objectivity (1999, 356). By contrast, *Political Liberalism* is concerned with the conditions necessary for stability (1996: xxxvii). Stability is only realizable if people with realistic reasoning capacities are willing to accept the principles that regulate society, but the Rawlsian notion of objectivity depicted here appeals to idealized reasoning capacities. Achieving reflective equilibrium requires more reflection than most can dedicate to it. For Rawls, the burdens of judgment are explicitly offered as a reason why we would not reach agreement on what is objective (Rawls 1996: 121). So, what conceptions people would accept in reflective equilibrium cannot be an adequate basis for stability. This does not deny that a political conception of justice has a kind of objectivity, but it is a limited kind of objectivity focused on what is reasonable. So it would be a mistake to interpret the discussion with Hart as a precursor to the arguments of *Political Liberalism*. The project of finding agreement among reasonable people marked by the burdens of judgment is quite different from the project of finding agreement among people that have achieved reflective equilibrium.

4.4

The final objection argues that it is too optimistic to think that there will be an alignment between fundamental conceptions in reflective equilibrium. For there to be any standard of correctness for moral judgments, there must be some conception of the person, some value, or some principle that is shared across persons in a justificatory community. The final objection claims that persons are too different to reach similar moral conclusions. We should not hope for shared fundamental conceptions, so this account cannot provide a standard of correctness for moral judgments.

In response, I emphasize how different it is to think that (a) actual persons will have shared fundamental conceptions, and that (b) persons who have achieved reflective equilibrium will share fundamental agreement. It is unlikely that any sizable proportion of people currently share the same fundamental conceptions, but it is much more plausible to think that people would share fundamental conceptions in reflective equilibrium. At that stage, inconsistencies will not be an issue, implicit commitments will have been scrutinized, and latent biases will have been exposed. So, much of what underlies disagreement between actual persons will have been washed away. At this stage, the similarities that people share as agents and because of their social atmosphere have a greater chance of explaining shared conceptions.

Even if it is optimistic to think that there will be fundamental conceptions that are shared between those in reflective equilibrium, it is not *too* optimistic. After all, this argument does not need to prove that there *is* a standard of correctness for moral statements. Instead, it needs to show *how we can conceive of moral objectivity*. In the end, the possibility of alignment in conceptions is all we need for most theoretical problems. We can only make sense of moral judgments, language, justification, and experience if we suppose a possible standard of correctness. There is a standard of correctness for moral judgments—and thereby moral objectivity—if there would be an intersubjective accord between fundamental conceptions in reflective equilibrium. Whether there actually is this accord is open to dispute. It seems to me that it should be. Rawls himself says that whether objective reasons actually obtain ‘can only be shown by the overall success over time of the shared practice of practical reasoning’ (1996: 119).

4.5

Stepping back from my argument, it seems that the reliance on optimism is the most significant contrast between the Rawlsian and Kantian conceptions of objectivity. The Rawlsian account shows how a standard of correctness is possible, but it does not prove that there is such a standard. Since we cannot know whether there would be shared fundamental conceptions in reflective equilibrium, we cannot prove that there is a standard for moral objectivity. Instead, we should go through the process of reflection, appeal to current deep convictions, and hope for the best. By contrast, the Kantian account aims to prove that there is moral objectivity through an account of agency. For me, this is a substantive appeal of the Kantian project. What could be a more worthwhile philosophical project than the vindication of moral objectivity? However, the requirements on an account of agency can seem quite demanding. For those who are skeptical about this Kantian project, the hope for objectivity that the Rawlsian project provides has an appeal of its own.

Even if we cannot prove that there would be an alignment of conceptions in wide reflective equilibrium, we can recognize that such an alignment would provide a standard of correctness. Moreover, it would be a standard of correctness that was consistent with our own deepest convictions. According to the Rawlsian account, we can make sense of moral objectivity in this way as we continue to seek

reflective equilibrium. According to this conception of objectivity, we should continue the process of moral theory and seek an agreement in our fundamental conceptions.

C. M. MELENOVSKY
SUFFOLK UNIVERSITY
cmelenovsky@suffolk.edu

References

- Daniels, Norman. (1980) 'Reflective Equilibrium and Archimedean Points'. *Canadian Journal of Philosophy* 10, 83–101.
- Enoch, David. (2006) 'Agency, Shmagency: Why Normativity Won't Come from What Is Constitutive of Action'. *Philosophical Review* 115, 169–98.
- Freeman, Samuel. (2007) *Rawls*. New York: Routledge.
- Korsgaard, Christine. (2009) *Self-Constitution: Agency, Identity, and Integrity*. Oxford: Oxford University Press.
- Laden, Anthony Simon. (2013) 'Constructivism as Rhetoric'. In Jon Mandle and David Reidy, (eds.), *A Companion to Rawls* (West Sussex: John Wiley & Sons), 57–72.
- Rawls, John. (1981) Letter to HLA Hart [April 22, 1981], Papers of John Rawls; Harvard Archives, HUM 48, Box 46, Folder 10.
- Rawls, John. (2000) *Lectures on the History of Moral Philosophy*. Cambridge, MA: Harvard University Press.
- Rawls, John. (1996) *Political Liberalism*. New York: Columbia University Press.
- Rawls, John. (1999a) *Collected Papers*. Cambridge, MA: Harvard University Press.
- Rawls, John. (1999b) *Theory of Justice*. Cambridge, MA: Harvard University Press.
- Ronzoni, Miriam. (2010) 'Constructivism and Practical Reason: On Intersubjectivity, Abstraction, and Judgment'. *Journal of Moral Philosophy* 7, 74–104.
- Scanlon, T. M. (2003) 'Rawls on Justification'. In Samuel Freeman (ed.), *Cambridge Companion to Rawls* (Cambridge: Cambridge University Press), 139–67.
- Street, Sharon. (2010) 'Constructivism about Reasons'. *Philosophy Compass* 5, 363–84.
- Quinn, Warren. (1987) 'Reflection and the Loss of Moral Knowledge: Williams on Objectivity'. *Philosophy and Public Affairs*, 16, 195–209.