

Constitutivism about Reasons: Autonomy and Understanding¹

Karl Schafer

Constitutivism about reasons begins with the idea what we have reason to believe or do is grounded in facts about our nature as acting, believing, and reasoning beings. Thus, constitutivism claims that our nature as rational agents takes explanatory priority over facts about our reasons. We have the reasons we do *because* we are rational agents of a certain sort.

Such a view has seemed attractive to many.² But in developing it, we immediately face a difficult question. Like anything else, I may be accurately thought of in many ways. And many of these descriptions involve an implicit conception of my proper activity or function. In other words, to use Thomson's terminology, I fall under a wide variety of different "goodness fixing kinds".³ Moreover, many of these kinds might reasonably be thought of as providing us with a conception of myself as a rational agent. And any of these might in principle be used as the foundation for a constitutivist account of my reasons. For example, we might begin here with a minimal conception of rational agency of the sort familiar from "Humean" accounts of rationality.⁴ Or we might begin with a more expansive conception of rationality of the sort characteristic of Kantian work on these issues.⁵ Or we might begin with the form of rational agency that is particular to *human beings* in the manner characteristic of neo-Aristotelian views.⁶

All of these have some plausibility, but we cannot say that all of them are relevant in the same way to what I have reason to do. For these descriptions often involve conflicting conceptions of my proper function or activity. And, as such, they are likely to generate competing accounts of what I have reason to do. Thus, in developing a constitutivist account, we need to determine *which* characterization of my nature grounds my reasons.

¹ ACKNOWLEDGEMENTS. Note about "Rationality as a Capacity for Understanding".

² This view is, of course, associated with the Aristotelian tradition. For recent forms of constitutivism see Foot(2003), Katsafanas(2013), Korsgaard(1996, 2009), Smith(2012), Thompson(2008), and Velleman(2009), amongst others. I'll focus primarily on Korsgaard's and Velleman's views, but much of what I say about them applies in some form to Katsafanas(2013). Smith's view is closer to my own in some ways - although there are many differences between them.

³ Thomson(2008). For a pluralistic view of agency, see Lavin(forthcoming). My view is like Lavin's in allowing for a variety of forms of agency - but unlike it in taking one very abstract characterization of these forms to have foundational significance for facts about reasons.

⁴ Street(2008), and (to some extent) Smith(2012, 2013) and Manne(forthcoming).

⁵ Bagnoli(2013), Engstrom(2013), Korsgaard(1996, 2009), Velleman(2000, 2009), and Schafer(2015a,b).

⁶ Foot(2003), Thompson(2008).

Nor it is obvious that non-normative inquiry into metaphysics can settle these questions. For example, while I can be described in many ways, perhaps only one of these accurately characterizes my fundamental essence as an acting thing. And perhaps it is this description that we should focus on when developing a constitutivist account of my reasons. But it is not obvious that a characterization of me - solely in virtue of being *metaphysically* fundamental - represents the level of description that determines what I have reason to do. And it might be that we can only know my essence as an acting thing *via* consideration of what my reasons for action are. For example, following Kant, perhaps we can become aware of our essence as agents only through a prior knowledge of what we ought to do. If so, then only through thinking about foundational questions in metaethics will I be able to know my essence as a rational agent.

For similar reasons, we shouldn't assume that the constitutivist *must* begin with the conception of agency or intentional action that is the focus of contemporary work in the philosophy of action.⁷ For example, the conception of rational agency that is relevant to constitutivism may involve more than the mere capacity to "act intentionally" in the ordinary English sense of these words. Indeed, as discussed below, I think that the proper starting point of the constitutivist project is best described - not as conception of agency *per se* - but rather as a general conception of *rationality*, of which rational agency is one form.

1. Reflection and the Concept of a Reason

So how should the constitutivist determine the description of rational agency that is relevant to her project? A tempting thought, made famous by Korsgaard's discussion in *The Sources of Normativity*, is that we can answer this question by considering the sort of creatures for whom questions about reasons arise in the first place.⁸ As Korsgaard somewhat metaphorically puts this point, we can think of terms like "reason" as picking out solutions to certain sorts of normative problems - problems that arise only for creatures of a particular sort. If so, then to understand what reasons are, we should begin by asking ourselves what it is to solve these problems *well*. And to understand this, for the constitutivist at least, we need to understand what it is to be such a creature.

Korsgaard describes such beings as follows:

Normative concepts exist because human beings have normative problems. And we have normative problems because we are self-conscious rational animals, capable of reflection about what we ought to believe and to do. That is why the normative question can be raised in the first place ... (46-47)

Korsgaard expands on this idea as follows:

⁷ Contrast Setiya(2009).

⁸ Korsgaard(1996). I won't consider how much Korsgaard's views have changed over the last two decades. But my sense is that she would regard her views as relatively constant at their core, even while her expression of them has shifted.

But the human mind *is* self-conscious in the sense that it is essentially reflective. ... And this sets us a problem no other animal has. It is the problem of the normative. For our capacity to turn our attention on to our own mental activities is also a capacity to distance ourselves from them, and to call them into question. I perceive, and I find myself with a powerful impulse to believe. But I back up and bring that impulse into view and then I have a certain distance. Now the impulse doesn't dominate me and now I have a problem. Shall I believe? Is this perception really a *reason* to believe? I desire and I find myself with a powerful impulse to act. But I back up and bring that impulse into view and then I have a certain distance. Now the impulse doesn't dominate me and now I have a problem. Shall I act? Is this desire really a *reason* to act? The reflective mind cannot settle for perception and desire, not just as such. It needs a *reason*. Otherwise, at least as long as it reflects, it cannot commit itself or go forward. (92-94)

Thus, for Korsgaard, the need for the concept of a reason arises because we are capable of self-conscious reflection, which allows us to take a step back from our perceptions, desires, and other attitudes and ask questions about whether and why we should accept those attitudes.

There is, I think, something right about this picture of how the need for the concept of a reason emerges: namely, that its natural home is in a certain sort of reflective deliberation.⁹ It is important to stress that to say this is not to make the over-intellectualist claim that the only manifestations of rationality involve conscious reflection or deliberation. Rather, the point here is about the concept of a reason. In particular, one can accept that the need for this *concept* emerges only in reflective contexts, while also claiming that the sort of *responsiveness to reasons* which is characteristic of rational thought is found in many less reflective contexts.¹⁰

Thus, I think that Korsgaard's starting point is far less intellectualist than it might seem, although of course it will be too intellectualist for some.¹¹ But there is also something surprising about Korsgaard's description of the sort of reflection at work here. As Korsgaard describes it, in engaging in this sort of reflection, we normally take a step back from our perceptions and desires and make *those perceptions and desires* the object of reflective inquiry. In other words, in such reflection, we ask, of the various elements of our state of mind, whether *these elements* stand up to reflective scrutiny.

⁹ Compare Burge(1986) and Smithies(2015).

¹⁰ For helpful discussion, see Katsafanas(2013).

¹¹ Indeed, I'm happy to acknowledge that there are senses of "rationality" that are much less closely connected with conscious reflection than the concepts I am exploring here. But I hope that the claim that there is a concept of "reason" or "rationality" that has something like this character, and which is philosophically significant, is not too implausible.

In this way, as a number of critics have noted, Korsgaard's account seems to mischaracterize the phenomenology typical of such reflective episodes.¹² Normally in such cases, far from turning inward, our attention remains focused outward on the world. In other words, in such reflection we remain focused on the question of what is true or what is to be done - and not on whether our *attitudes* are correct or incorrect.¹³

This may seem a minor point, but Korsgaard's understanding of reflection resonates through her view. For Korsgaard, the aim that is constitutive of agency is to unify one's self - or, in other words, to achieve genuine autonomous agency. This conception of the aim of agency makes perfect sense if the sort of reflection at issue here is paradigmatically reflection on the status and relations of our own attitudes. For in that case, what will become salient to us through such reflection are the ways in which our attitudes and actions are in conflict with one another. Thus, such reflection will generate a consciousness of our own lack of unity. Given this, it is natural to think of the primary challenge that such reflection poses in terms of overcoming this disunity, in very much the manner Korsgaard describes.

But if we think of rational reflection in a more object-directed or world-directed way, this will not seem so obvious. For the primary task that such object-directed reflection poses is the task of knowing or understanding what is the case or what is to be done. And, at least at first glance, this is not obviously the same as the task of unifying one's own mental states or attitudes as such.

In this way, Korsgaard's conception of the manner in which reflection gives rise to questions about reasons makes it natural for her to focus on the task of achieving a certain sort of subjective or agential unity. But if we conceive of such reflection in a more object-directed fashion, our attention will focus somewhere else: namely, the task of achieving knowledge or understanding of what is and what is to be done. In this way, one's understanding of the sort of reflection that is relevant here naturally informs one's view of the aim of rational agency.¹⁴

Moreover, once we conceive of rational agency in terms of the aim of achieving an understanding of what is to be done, it becomes natural to think of rational agency as

¹² Compare Blackburn(2000). For related worries about more Humean views, see Schroeder(2009). For Korsgaard's response, see Korsgaard(2008).

¹³ There is a grammatical point that might seem to count against this - namely, that reasons are often described as reasons *for belief* or *for desire* or *for intention*. This does indicate that the need for the concept of a normative reason only emerges for creatures that are capable of distinguishing appearance from reality. Thus, in order to phrase questions about reasons we often need to advert to language that allows us to draw this distinction. But this does not mean that such reflection is primarily concerned with these attitudes.

¹⁴ Similar issues arise with respect to Velleman's form of constitutivism, but there things are complicated by Velleman's characterization of the sort of autonomy that is distinctive of rational agency in terms of a certain sort of understanding. Nonetheless Velleman also characterizes this sort of understanding as an understanding *of oneself* as an agent. Thus, in the end, he shares with Korsgaard a picture of the constitutive aim of rational agency on which this aim is directed at oneself. I return to this below.

one manifestation of a more general capacity for both theoretical and practical understanding. I will say more about this shortly. But for now, I just want to note that this makes it attractive to think of the constitutivist project - not as beginning with a notion of *agency* in particular - but rather, as beginning with a general notion of *rationality*, of which rational agency is one form.

That having been said, I do *not* mean to deny that there is a constitutive connection between rationality and autonomy or subjective unity. Indeed, to jump ahead for a moment, on my view, rationality does have a constitutive connection with a certain sort of autonomy. But this is *because* this sort of autonomy is the "subject-directed" analogue of the sort of "object-directed" understanding at which rationality primarily aims. Thus, in the end, I agree with much of what Korsgaard says about these issues - albeit in a somewhat different context than she would accept.

My aim in this essay is to explore this sort of "understanding-first" constitutivism. To do so, I'll proceed as follows. First, I'll discuss how the nature of rational reflection supports a conception of rationality as the capacity for a certain sort of understanding. Then, I'll explain why this general capacity has both a theoretical and a practical manifestation, by discussing the theoretical and practical forms that such understanding can take. In this section, I'll also address some worries about the circularity of this account. With that in mind, I'll turn to the forms of reasoning that are essential to rationality in this sense. In particular, I'll sketch how this account of rationality might be used to explain (i) the rationality of forms of explanatory reasoning and (ii) a responsiveness to the views of others. Finally, I will return to the relationship between understanding-first and autonomy-first constitutivism.

2. Rationality and Understanding

Let's begin with the conception of rationality that sits at the heart of this form of constitutivism. Above I followed Korsgaard by taking the sort of reflection within which questions about reasons arises as a *Leitfaden* for thinking about the kind of rationality that is constitutively connected with reasons. The hope here is that, by doing so, we can isolate the conception of rationality of interest to the constitutivist by considering the capacities and activities involved in bringing such reflection to a successful conclusion. For if the "reasons problem" arises in course of such reflection, then to understand the proper "solution" to this problem, we must understand what it is to bring such reflection successfully to an end.

Thus, we can think of rationality in the sense of interest to us in terms of the capacity that is required to *successfully* engage in this sort of reflection.¹⁵ To think of rationality in this way is to adopt a broadly virtue-theoretical understanding of rationality in terms of the proper functioning of a *capacity*. Such capacities are characterized by a distinctive aim or function.¹⁶ As a result, rationality will be associated with certain standards of "goodness" which can be derived from this characteristic activity or function. In this way, the idea of rationality as a capacity

¹⁵ Of course, "rationality" is used in contemporary English to pick out a variety of properties and states. In discussing these issues, I don't mean to provide a comprehensive account of rationality in all these various senses.

¹⁶ For similar ideas, see Svavarsdóttir(2008) and Wedgwood(2015).

generates both a "capacity conception" of rationality and an "evaluative conception" of what it is for this capacity to function *well*.¹⁷

But what is it for *this* capacity to function well? As we just discussed, reflection on the phenomenology of rational reflection suggests that this function is primarily concerned with achieving theoretical or practical understanding, as opposed to autonomy or self-unification. But we need to be cautious in inferring facts about the constitutive function of rational reflection from its phenomenology.¹⁸ After all, it might well be that the nature of rational reflection guides our attention away from its constitutive function.

Nonetheless, it is important to remember that the phenomenology of rational reflection is not merely a matter of where such reflection directs our attention. Rather, the phenomenology of rational reflection involves an implicit understanding of *its point*. Of course, this aspect of the phenomenology of reflection might be misleading. But nonetheless it is reasonable to treat it as a defeasible starting point in considering these questions.

To flesh out this point, consider what it is for this sort of reflection to come to a fully successful conclusion. Such reflection is primarily concerned with two questions, one theoretical and one practical: namely, the question of *whether P*, and the question of *whether to A*. It is these sorts of questions that theoretical and practical reflection seems to aim at answering, at least in the first instance.

This suggests an initial answer to our question. For given that rational reflection is concerned with such questions, isn't it obvious that it will reach a successful conclusion just in case it arrives at the *correct* answers to them? Given this, it seems natural to think that the function or aim of rational reflection is to help us reach correct conclusions about these questions. More precisely, on this view, the function of theoretical rationality would be to produce true beliefs. And the function of practical rationality would be to produce correct intentions or actions, however this notion of "correctness" is understood.

Especially in the practical domain, such a view might be made precise in a variety of ways - corresponding to different accounts of the "practical correctness" that corresponds to "truth" in the case of beliefs. But even if we focus on the case of theoretical reflection, where things are simplest, I think this simple account provides us with an incomplete picture of the function of rational reflection.

Of course, theoretical reflection or inquiry will only reach a fully successful conclusion if it produces accurate representation of what is the case.¹⁹ The question is whether this is *all* that a fully successful conclusion to such reflection requires. I think it is clear that it is not. After all, we may form a true belief for all sorts of

¹⁷ This conception of rationality has a noble historical pedigree. But it is not the only - or even the most dominant - account of rationality today. For further discussion, see Schafer(forthcoming).

¹⁸ Korsgaard(1996, 2009), Velleman(2009).

¹⁹ Hieronymi(2005), Shah(2003), Shah and Velleman(2005).

reasons and in all sorts of ways. And such beliefs can, and often will, fall sort of valuable epistemic statuses such as knowledge or understanding. When they do, it seems that, while the rational reflection in question has come to a *correct* conclusion, it is less than *fully successful*.

For example, philosophers often speak of theoretical inquiry as successfully terminated by *knowledge that*. This already involves a more demanding conception of the aim of rational reflection than truth. But, at least as it is often understood, it also leaves something essential out of the picture. For the fully successful conclusion of rational reflection involves more than mere propositional knowledge, at least if this is understood in a piecemeal fashion. In particular, while knowledge that P may settle the question *whether P* in some sense, it need not terminate inquiry *into P* more generally. For example, even if I know that P, it may still be the case that I have no idea why or how P is true. Thus, I may not be able to *make sense* of P. If this is the case, I will have good reason to prolong my inquiry into P - namely, that I fall short of full understanding of it.²⁰

In other words, inquiry is concerned, not just with knowing *that*, but also with knowing how, knowing why, knowing which, knowing what, and so on. For example, if I am unable to cite any reasons why P is true, I will be unable to explain why my view is the correct one. And if I am unable to use these reasons as my grounds for believing P, I will have failed to fully integrate this belief into my general understanding of the world.²¹ In these cases, there is a clear sense in which rational reflection about P is incomplete. For while I may know that P, I will be unable to satisfy the request for reasons that arises in the context of rational reflection.

Thus, while there is a close connection between understanding and knowledge, understanding involves more than mere piecemeal knowledge that. Rather, in the sense I am interested in, understanding is something more like *a properly structured knowledge of that, how, why, which, etc.*²² One consequence of this is that, while knowledge is often thought of as something one either has or lacks, understanding is the sort of thing that generally comes in degrees. For one's level of understanding will increase with the degree to which one has the required pieces of knowledge in the right relations.

This distinction between piecemeal propositional knowledge and genuine understanding is a familiar one. For example, understanding biology involves more than knowing a variety of disconnected biological facts. It also requires that I can make sense of these facts. And this requires, not just various pieces of knowledge that, but also knowledge how and why and so forth. For this provides one's understanding of biology with the structure that is characteristic of understanding.

²⁰ For a similar arguments, see Kvanvig(2003) and Pritchard(2010). For some of the other recent literature on understanding, see Elgin(1993), Grimm(2010), Hills(2015), and Zagzebski(2001).

²¹ For the distinction between merely being able to cite reasons and being able to use them as reasons in this context, see McGrath(2010) and Hills(2015).

²² See Sliwa(2015).

Moreover, this phenomenon is easy to locate in the practical domain as well. For example, practical understanding of what is to be done requires more than merely intending to do the right thing. It also requires that one can make sense of why this is to be done and how. Indeed, much of ordinary instrumental reasoning can be understood as aiming at a deeper and more systematic understanding of the thing to do. In this way, even very basic forms of practical reasoning can be seen to contributing to practical understanding - for example, by giving the agent knowledge that she is doing A in order to do B.

This is closely connected with the idea that it is this sort of reflection that gives rise to the need for the concept of a reason. For the function of the concept of a reason is not simply to improve the *correctness* of our answers to such questions. Rather, this concept also allows us to grasp the explanatory connections between different questions and answers. Thus, the fact that rational reflection gives rise to this concept indicates that the fully successful conclusion of such reflection involves more than piecemeal knowledge that. Rather, such reflection is only fully successful when we possess the sort of understanding that allows us to *make sense* of our answers to the questions we are considering.

In this way, the nature of rational reflection supports the conclusion that such reflection is only fully successful insofar as it leads to genuine understanding. And given its tight connection with the conditions under which rational reflection is fully successful, I believe this conception of understanding is ideally suited to characterize the function of rationality in the sense of interest to us here. Moreover, as I explain below, this conception of rationality should also be attractive to a constitutivist because it provides us with extra resources for making sense of reasons that are difficult to explain on more restricted conceptions of rationality. But before addressing this issue, I want to say a bit more about this conception of rationality and the correlated notion of understanding - focusing on its ability to provide us with a unified understanding of theoretical and practical rationality as two manifestations of a general capacity for understanding.

3. Theoretical and Practical Rationality

To do so, we need to consider what distinguishes theoretical from practical rationality. On this view, we can understand these forms of rationality in terms of the general capacity for rationality insofar as it is manifested by theoretical or practical thought. So the requirements of epistemic and practical rationality *in particular* will follow from this general conception of rationality when it is combined with the *particular* functions of theoretical or practical thought. In other words, both theoretical and practical thought have the function characteristic of rational thought in general - namely, that of generating understanding. But theoretical and practical thought are distinguished from by the *particular* sort of understanding it is their function to produce.²³

But how should we distinguish these forms of understanding? There are at least three main answers to this question. First, there is the idea that theoretical understanding is distinguished from practical understanding in terms of the different *standards of*

²³ Engstrom(2013) and James(2012). Contrast Velleman(2009).

correctness that are constitutive of the understanding in question.²⁴ Second, there is the idea that theoretical and practical reasoning aim to answer importantly different sorts of *questions*.²⁵ And third, there is the idea that basic theoretical/practical distinction involves the explanatory relationship between one's understanding and what is understood - so that theoretical understanding is understanding that is grounded in its object's existence in some sense, while practical understanding has the opposite relationship to its object.

I don't think these should necessarily be seen as incompatible answers to our question. Rather, a version of all three of these claims may be true.²⁶ But the important point for us is simply that, whichever of these views we adopt, we can understand the difference between theoretical and practical rationality in terms of a more basic distinction between these forms of understanding.

To summarize, while remaining as neutral as possible between these three formulations, let's say that the distinctive function of theoretical reason is to understand *the way things are*. Correspondingly, the distinctive function of practical reason is to understand *what to do* - and thereby to shape the world through our understanding.²⁷ In making these claims, it is important not to confuse the distinctive function of a form of thought with its standard of correctness.²⁸ For example, while the *characteristic function* of practical thought seems to involve effective action in some way, the *standard of correctness* for it will be the practical analogue of the *true* - namely, what is to be done or what ought to be done.²⁹

This has important implications for the relationship between my view of practical rationality and Velleman's form of constitutivism. Although his view does conceive of practical rationality in terms of a certain sort of autonomy, Velleman also takes this sort of autonomy to be closely related to a sort of understanding. Thus, like me, Velleman takes the function of practical rationality to lie in a certain sort of understanding. The difference between us lies in how we understand the relevant sort of understanding. For me, the understanding that is relevant to practical rationality is just the practical analogue of theoretical understanding. So, just as theoretical rationality aims to understand what is, practical rationality aims to understand what is to be done. For Velleman, on the other hand, practical rationality aims at a certain sort of folk-psychological understanding of *oneself* as agent. Thus, on his view, the aim of practical reasoning and reflection is to produce actions that are intelligible to the agent

²⁴ Wedgwood(2007), Velleman(2000), Way and McHugh(forthcoming).

²⁵ Hieronymi(2005), Shah and Velleman(2005).

²⁶ For example, Kant seems to draw this distinction in all these ways. See Schafer(2016).

²⁷ None of this is meant to suggest that it is possible to neatly factor our understanding into a theoretical and a practical component. On the contrary, it is very plausible that much of our theoretical and practical understanding of things is intertwined in various ways.

²⁸ Thus, talk of the "direction of fit" of such states is quite dangerous in the present context, since it tends to paper over this distinction. Compare Frost(2015).

²⁹ For more discussion, see Schafer(2013).

given his folk-psychological understanding of himself.³⁰ In this way, for Velleman, the understanding that is constitutive of practical rationality is fundamentally self-directed and ultimately theoretical in character.

Which of these accounts should we prefer? Velleman's defense of his view against objections is nuanced, so I do not mean to present a knockdown argument against it. But for reasons already mentioned, it seems to me that my view is closer to the ordinary first-personal understanding of the *point* of practical reasoning. For when I am reasoning about what to do, it does not seem to me that I am engaged in an activity that aims at achieving a better folk-psychological understanding of myself.

Of course, such considerations are hardly decisive on their own. But they provide us with a *prima facie* case in favor of my view. And I think that, at least for a Kantian, there are other reasons to prefer my account. Velleman's view treats the sort of understanding that is characteristic of practical rationality as a subspecies of theoretical understanding - albeit one with a self-fulfilling character. This, contrary to Kant, gives theoretical rationality a sort of priority over practical rationality. We can avoid this by viewing practical rationality as providing us with a sort of practical understanding that cannot be reduced to a special instance of theoretical understanding. Thus, if our goal is to understand theoretical and practical rationality as two *equally basic* manifestations of a general capacity for understanding, I think we can do better here. In other words, in this way, Velleman doesn't seem to me to do full justice to the possibility of *genuinely practical* reason. This, I think, would be Kant's main concern about Velleman's account.

But what of Velleman's motivations for characterizing practical rationality in the manner he does? In rejecting the sort of view I've been defending, Velleman is motivated by the thought that such a view make it impossible for the constitutivist to explain reasons or values in terms of the nature of rationality or agency. For example, he writes:

Whereas most philosophers think that the intelligibility of an action derives from its appropriateness or rightness or goodness, real or perceived, I am going to argue that appropriateness or rightness or goodness derives from [folk psychological] intelligibility - which cannot then derive from such normative considerations on pain of circularity. My reason for reversing the order of explanation is that I can thereby account for what is less clear, normativity, in terms of what more clear, psychological explanation... (Velleman(2009), 27)

This is an important point. But, for two reasons, I don't think it counts against understanding-first constitutivism, once it is properly understood. First, while my account does characterize the function of practical rationality in terms of a distinctive sort of practical understanding, it need not conceive of practical understanding in explicitly normative terms. Rather, it can conceive of what is distinctive of practical understanding in terms of the relationship between such understanding and its object, or in terms of the sorts of questions it aims to answer, or simply as an understanding

³⁰ Velleman(2009): "The relevant notion of 'making sense' is not normative: it's not about what the character ought to do. It's the notion of what can be understood in terms of the character's attributes and attitudes under the circumstances." (13)

of *what is to be done*. The last of these possibilities is particularly important. For it indicates that we need not begin by characterizing practical understanding as involving an understanding of the "good" or what "ought to be". Rather, we can also characterize it in broadly *imperative* terms.³¹ By doing so, we can capture what is distinctive about practical understanding without characterizing practical understanding as an understanding of normative facts.

But, and this is the second point, even if we do characterize practical understanding as a sort of understanding of such facts, this will not necessarily make the attempt to explain facts about what ought to be done in terms of prior facts about rationality circular. For there is nothing circular in first characterizing practical rationality in terms of certain thoughts involving the *concept or mode of presentation* OUGHT or GOOD, and then using this characterization of rationality to explain the reference and the truth-conditions of these thoughts. This will be possible, for instance, provided that it is possible to characterize the nature of ought-thoughts (on the level of sense) by specifying the conditions for understanding them, and more specifically, the conditions for possessing the normative concepts they deploy. Such a characterization needn't involve a characterization of the reference of these concepts. Instead, we might characterize these possession conditions via a characteristic role the concept plays in cognition and action. We can then use this characterization of the possession conditions of the concept as input into the determination of the reference and truth conditions of those ought thoughts. Such an approach is by no means trivial, but it is familiar enough from the discussion of concepts whose truth-conditions are determined by their proper use.³²

Thus, these circularity worries are no reason to reject a characterization of practical rationality as aiming at an understanding of what is to be done. Of course, this does not rule out the possibility that fully realizing this function requires something like the sort of self-understanding as agents that Velleman describes. Indeed, it seems to me that something like this is the case. But acknowledging this does not require us to reject a classical picture of practical reasoning as aiming at understanding what is to be done.

4. The Requirements of Understanding

The foregoing provides us with a general strategy for the constitutivist. In the practical domain, this begins with a conception of rationality and its success conditions, characterized in terms of understanding what to do. The constitutivist strategy is, then, to derive facts about what we ought to do and the reasons we have from general constraints on understanding. To demonstrate the power of this approach, I will briefly discuss two central ways in which the nature of understanding places non-trivial constraints on the nature of good reasoning, with the hope that this will help to demonstrate the power of this approach.

³¹ See Schafer(2013). This is not far from the understanding of practical thought developed by expressivists like Gibbard(1992, 2003) and Blackburn(2000).

³² For an example of this sort of strategy in a realist context, see Wedgwood(2007). For some relevant background, see Peacocke(1992).

a. Explanatory Reasoning

To do so, let's begin with the significance of explanatory reasoning for rationality, conceived of as a capacity for understanding.

As noted above, theoretical understanding of some domain of facts requires more than a scattered and piecemeal collection of true beliefs. Rather, understanding requires a grasp of the *explanatory connections* within that domain. For example, to understand physics, I must have some grasp of the why particular physical facts follow from certain physical laws - and in doing so I must grasp the explanatory connections between the two. For only then will I be in a position to articulate why some physical fact is true.³³ And similarly, to really understand what is to be done, I must have some implicit grasp of *why* particular actions are to be done. For example, I may grasp that some action A is to be done as a way of doing B. So here too, my intentions will only constitute genuine practical understanding if they possess this sort of explanatory unity.

But in order for my beliefs or intentions to have this sort of unity, they must be connected together in certain ways. That is, they must form a network of inferences. And the inferences in question must go beyond encoding merely logical relations. Rather, in order to capture the explanatory links we are interested in, these inferences will often have to involve explanatory forms of reasoning like inference to the best explanation.³⁴

For instance, part of an explanatorily unified set of beliefs is a grasp of how explanatory connections are relevant to ways in which these beliefs *ground* each other. For example, if I grasp an explanatory connection between events of type A (explanans) and events of type B (explanandum), then I will be disposed to make a defeasible inference from the existence of an A-event to the existence of a B-event. That is, I will be prepared to see belief in one as a potential ground of belief in the other. And similarly, if I believe that the best explanation of events of type B is normally the presence of an event of type A, I will be disposed to make a defeasible inference from the existence of a B-event to the existence of an A-event. In this way, part of having an explanatorily unified set of beliefs is ability to connect these beliefs together via various inferential relations. And these relations go beyond deductive inference to include various forms of explanatory or ampliative inference. Thus, explanatorily unified understanding requires that one's beliefs be taken to ground each other in a variety of explanatorily significant ways.

³³ As Grimm(2010) writes: "understanding is directed at a complex of some kind - in particular, at a complex with parts or elements that depend upon, and relate to, one another, and that the mind grasps or apprehends when it understands."

³⁴ For evidence of this, we might consider cases in which one *knows* that something happens *because* of something else, but does not *grasp how* the one event could ground the other. For example, knowledge that P occurred because of Q can be acquired through ordinary testimony - but this is often insufficient for one to fully understand this connection. See Hills(2015).

If this is right, then genuine understanding requires that our attitudes are connected together by explanatory forms of inference. For it is only insofar as this is true that we can truly be said to have *internalized* the explanatory relations that are relevant to understanding. In this way, the acceptance of such patterns of inference is partially constitutive of understanding. Thus, the capacity to make such inferences is partially constitutive of the capacity for understanding. And, of course, making such inferences is also an excellent means of improving our understanding.

Now, there may be a variety of inferential methods that are capable of playing this role. So there may not be any *particular* form of explanatory inference that is required here. But what is so constitutive is the use of some sort of inference that gives preference to more explanatory sets of beliefs and intentions. Thus, if the function of rationality is understanding, we can see why we are rationally required to use *some* form of reasoning that privileges better explanations over worse.³⁵ And similarly, we can see why the use of such methods of reasoning is generally more rational than forms of reasoning that do not have this character. In this way, the idea that rationality is a capacity for understanding is well placed to explain the special rational status of explanatory inference.³⁶

b. Understanding and Intersubjectivity

This is important because explanatory inferences of this sort play such a large role in our ordinary understanding of good reasoning *qua* a proper responsiveness to reasons.³⁷ Thus, a plausible form of constitutivism should begin with a conception of rationality that is capable of explaining why such inferences have a privileged rational status - in both the theoretical and practical domains.

Something similar is true of what we might call the "intersubjective dimensions" of rationality. For example, one way of being properly responsive to reasons for belief is to listen to the testimony of others. And one way to be properly responsive to reasons for action is to be sensitive to the concerns of others. Thus, a plausible form of constitutivism ought to begin with a conception of rationality that can explain the rationality of these patterns of thought.

³⁵ I discuss this in greater detail in Schafer(forthcoming).

³⁶ These conclusions follow, even though understanding - like knowledge - is factive and so requires correctness. Given this, it is true that a pattern of inference will only tend to generate (say) theoretical understanding insofar as it is reliable. So these arguments provide us with no guarantee that use of such methods will be provide us with understanding. But nonetheless, these arguments show that our *only hope* for achieving understanding is to use such methods. Thus, while use of these methods is not the only necessary condition on achieving the aim of rationality, it is one such condition. And this explains why these methods should have a privileged rational status - at least insofar as a we lack positive evidence of their unreliability.

³⁷ This is obvious in the theoretical sphere. And much the same is true of practical reasoning. For example, we often reject some desire as misleading about what is worth doing because it conflicts with our overall understanding of what is to be done - even if it does not conflict with any other felt desire. Thus, just as inference to the best explanation sometimes leads us to reject the way things seem to us in perceptual experience, similar considerations can lead us to call into question our felt desires in the practical domain. See Schafer(2013).

Fortunately, just as there is a connection between understanding and explanation, there also seems to be a connection between understanding and the ability to transcend our own particular subjective point of view. Thus, the idea that rationality is a capacity for understanding can also help to explain the special rational status of being responsive to what others believe or will.

There are a number of ways of approaching this point. For example, it seems that ideal understanding requires that one's views and reasoning be potentially shareable. As Hills puts it, in order to understand that p is true because of q , you must have certain capabilities - including capabilities that involve communication with others:

If you understand why p (and q is why p), then you judge that p and that q is why p and in the right sort of circumstances you can successfully: (i) follow an explanation of why p given by someone else (ii) explain why p in your own words (iii) draw the conclusion that p (or that probably p) from the information that q (iv) draw the conclusion that p' (or that probably p') from the information that q' (where p' and q' are similar to but not identical to p and q) (v) given the information that p , give the right explanation, q ; (vi) given the information that p' , give the right explanation, q' .³⁸

To have these capacities, we must reason in ways that are sensitive to whether other rational agents could share our reasoning. As a result, someone who aims at understanding will have priors that are "biased" towards intersubjective agreement. And this will make them responsive to the views of other rational subjects, insofar as they are also engaged in the project of developing shared understanding. So although testimony and deference to others are not particularly effective ways of gaining genuine understanding, someone who is concerned with understanding will tend to be more responsive to testimony than they otherwise would be. In short, as Manne puts it, understanding seems to be the sort of thing that it is not limited to any particular individual. Rather, it is the sort of thing that should be achievable, not just by reasoning alone, but also by *reasoning with other rational subjects*.

Obviously, this leaves many questions unanswered about exactly what sort of reasoning is required here. But what I want to stress here is the *prima facie* plausibility of the idea that conceiving of rationality as the capacity for understanding can help to explain the rational significance of reasoning in an intersubjectively acceptable fashion.³⁹ Of course, this is at most the beginning of an account. But hopefully it will give the reader some cause for cautious optimism about this project.

5. Understanding, Rationality and Reasons

³⁸ Hills(2009). For related ideas, see Manne(forthcoming), Markovits(2013), and Smith(2012).

³⁹ Of course, the degree to which this ideal is achievable will vary. Indeed, it is plausible that there is a tension between the aim of achieving this sort of intersubjective understanding and the aim of doing justice in that understanding to the *diversity* of subjective points of view. It is this tension, for example, that makes it plausible that our practical reasons involve both agent-neutral reasons and agent-relative reasons in some form. See Nagel(1986) and Korsgaard(2008).

Given all this, to understand something is not just a matter of getting it right. It also requires a systematic and explanatorily powerful understanding of *why* things are as one takes them to be. In other words, it involves being in a position to give reasons for those thoughts - reasons that other rational subjects (at least in principle) can appreciate as valid.

This is not meant as an exhaustive characterization of the forms of thought that constitute ideal understanding. For example, I have said little here about the role of experience or emotions or sensible desires, or of empathy, or of forms of interpretation that fit awkwardly with the model of natural scientific understanding. Nor have I discussed other forms of "anti-Cartesianism" about understanding - such as the idea that understanding might be "extended" to include, not just narrow mental states, but also bodily needs and capacities of various kinds.⁴⁰ But I hope the two cases discussed above have helped to indicate how a focus on understanding can expand our sense of what rationality requires.

I also hope this makes it more plausible that rationality in this sense has a constitutive connection with what we have reason to believe and do. After all, normative reasons just are considerations that play a justifying role. And while this point is not stressed enough, this justifying role involves more than merely "counting in favor" of something in a minimal sense.⁴¹ Rather, to play this role, reasons must tap into some more general understanding of why something is to be believed or done - one that can be appreciated by other rational beings.⁴² Thus, the sort of *systematic* and *intersubjective* justification involved in understanding also seems to be essential to our understanding of what reasons *are*. This makes such a conception of rationality a natural place for the constitutivist about reasons to begin.

At the same time, the constraints that follow from the nature of understanding and its implications for rationality will often be *very* abstract. For example, these constraints will demand that one reason in ways that give *some* sort of priority to better explanations over worse. But this sort of reasoning might take many forms - and which of these forms is most appropriate for some agent might depend on their particular psychological capacities and cultural context. In this way, as Velleman nicely stresses, the sort of rationality at issue here will be multiply realizable. And the forms it takes in one culture may be different from the forms it takes in another.⁴³ Thus, what a particular individual has reason to do will be a function of these general rational constraints *as best realized in that individual's particular context*. But at the same time, this does not mean that there are no interesting constraints that apply to all rational agents as such. For the nature of understanding, while very abstract, does impose meaningful constraints of this sort - even if there are many ways for individuals and cultures to be true to them.

6. Autonomy and Understanding

⁴⁰ See Chalmers and Clark(2000). For discussion of bodily needs see Manne(forthcoming).

⁴¹ Compare Scanlon(2013) and Parfit(2011).

⁴² Korsgaard(2009).

⁴³ Velleman(2009).

With this in mind, let's return to the relationship between understanding-first constitutivism and autonomy-first constitutivism. As should be plain, the structure of these two views is very similar. Both views begin with a certain characterization of my nature as a rational agent, which carries with it a conception of my proper function or activity. And both use this function to derive further standards for action or thought via uncovering necessary conditions on autonomy or understanding, standards that in turn are used to explain what I have reason to believe or do.

But I think these forms of constitutivism are even more closely connected than this. One of the striking things about our discussion of rationality and understanding is that the constitutive standards that follow from this conception of rationality are very similar to the standards that many autonomy-first constitutivists take to follow from the connection between agency and autonomy. For example, in a similar fashion, Korsgaard has argued that genuine autonomy requires both (i) a certain sort of systematicity in one's will and (ii) a sort of transcendence of one's personal perspective.⁴⁴

From a Kantian perspective, the convergence of these views should actually be unsurprising. For it is a familiar Kantian point that achieving a unified perspective on the *objects* of thought requires unifying the *thoughts* by which one represents those objects. If this is right, then a certain sort of "subjective unity" will be a necessary condition of the "objective unity" that is characteristic of understanding. And something like this "subjective unity" seems to be what many contemporary autonomy-first constitutivists have in mind when they discuss autonomy or self-unification.⁴⁵ In this way, the understanding-first constitutivist may be able to derive the claims of the autonomy-first constitutivist as "subject-directed" corollaries of their "object-directed" claims about rationality and understanding.

Nonetheless, I think constitutivism about reasons is most compelling when it begins with understanding as opposed to autonomy. In part, this is because I find the idea that understanding is essentially connected with explanation and intersubjectivity somewhat easier to grasp than the idea that autonomy has such connections. Thus, constitutivist arguments seems to me most compelling when they proceed from understanding as opposed to autonomy.

But I also think that describing the constitutive aim or function of rational thought in terms of understanding the *objects* of thought - as opposed to the autonomy or unity of the thoughts themselves - better captures what matters to us from a first-person point of view. After all, in both theoretical and practical reasoning, our primary concern is not the degree of unity possessed by our thoughts themselves. Rather, our concern is with the *objects* we are thinking about. Thus, by characterizing the basic activity of rational thought in terms of understanding, we are truer to the "object-directed" character of rational thought.

⁴⁴ Korsgaard(2009).

⁴⁵ This helps to explain why self-consciousness is important for understanding - at least in the Kantian sense of self-consciousness in which an attitude is self-conscious just in case it is part of the relevant "subjective unity".

This is relevant to the worry that this form of constitutivism will do poorly at establishing the "first-person authority" of the relevant constitutive standards. Given understanding-first constitutivism, we must comply with the standards constitutive of rationality, on pain of a failure to understand. Similarly, given autonomy-first forms of constitutivism, we must comply with the standards constitutive of agency, on pain of ceasing to be an agent in the full sense of this term. In both cases, it is important to stress that this does not make it *impossible* to have doubts about the constitutive standards in question. It only indicates the costs associated with these doubts. Thus, from a structural perspective, these views involve similar forms of "inescapability". The only question is which captures a deeper sort of first-person inescapability.⁴⁶

And here it seems to me that the understanding-first constitutivist has the advantage over autonomy-first constitutivist. Quite simply, the "first-person authority" of understanding normally seems to me to be greater than the "first-person authority" of autonomy. Some evidence of this can, I think, be seen from the reactions non-Kantians often have to the significance that views like Korsgaard's attach to autonomy or self-unification.⁴⁷ But this idea gains support from the idea that fully successful inquiry into a theoretical or practical question ends with understanding. If that is right, then to appreciate the connection between the nature of inquiry and understanding, we need only to consider what it is for inquiry to succeed.

Of course, it may well be that successful inquiry *also* terminates in a state of mind that is autonomous in some sense. But it is the connection between inquiry and understanding that seems most straightforward from a first-person point of view. Thus, if we're looking for a constitutivism that provides us with "first-person inescapability", one that begins with understanding may be best placed to do so.

⁴⁶ Enoch(2006.2011), Tiffany(2012).

⁴⁷ See Gibbard(1999).

- Bagnoli, Carla. (2013) "Constructivism about Practical Knowledge." C. Bagnoli (ed.) *Constructivism in Ethics*. Cambridge: Cambridge University Press.
- (2011) "Constructivism in Metaethics." *Stanford Encyclopedia of Philosophy*.
- Blackburn, Simon. (2000) *Ruling Passions*. Oxford: Oxford University Press.
- Burge, Tyler. (1986) "Intellectual Norms and Foundations of Mind". *Journal of Philosophy*, 83(12): 697-720.
- Chalmers, David and Clark, Andy. (1998) "The Extended Mind." *Analysis* 58 (1): 7-19.
- Engstrom, Stephen. (2013) "Constructivism and Practical Knowledge." in C. Bagnoli (ed.) *Constructivism in Ethics*. Cambridge: Cambridge University Press.
- Elgin, Catherine. (1993) "Understanding: Art and Science". *Synthese* 95 (1): 196-208.
- Enoch, David. (2006) "Agency, Shmagency: Why Normativity Won't Come from What Is Constitutive of Action." *Philosophical Review* 115: 169–98.
- (2011) "Shmagency Revisited." in Michael Brady (ed.) *New Waves in Metaethics*. New York: Palgrave Macmillan.
- Foot, Philippa. (2003) *Natural Goodness*. Oxford: Oxford University Press.
- Frost, Kim. (2015) "The Very Idea of Direction of Fit." *Philosophical Review*.
- Gibbard, Allan. (1992) *Wise Choices, Apt Feelings*. Cambridge: Harvard University Press.
- (1999) "Morality as Consistency in Living: Korsgaard's Kantian Lectures." *Ethics* 110(1): 140-164.
- (2003) *Thinking How to Live*. Cambridge: Harvard University Press.
- Gopnick, Alison. (1999) *The Scientist in the Crib*. New York: William Morris.
- Grimm, Stephen. (2010) "Understanding." in D. Pritchard and S. Berneker (eds.) *The Routledge Companion to Epistemology*. London: Routledge.
- (2012). "The Value of Understanding" *Philosophy Compass* 7:2: 103-117.
- Hussain, Nadeem J.Z. (2012), "A Problem for Ambitious Metanormative Constructivism," in J. Lenman and Y. Shemmer (eds.), *Constructivism in Practical Philosophy*. Oxford: Oxford University Press.
- Hieronymi, Pamela. (2005) "The Wrong Kind of Reason." *Journal of Philosophy* 102 (9): 437-457.
- Hills, Alison. (2009) "Moral Testimony and Moral Epistemology." *Ethics* 120 (1): 94-127.
- (2015) "Understanding Why." *Noûs* 49(2).
- James, Aaron. (2012) "Constructing Protagorean Objectivity." in J. Lenman & Y. Shemmer (eds.), *Constructivism in Practical Philosophy*. Oxford: Oxford University Press.

- Katsafanas, Paul. (2013) *Agency and the Foundations of Ethics: Nietzschean Constitutivism*. Oxford: Oxford University Press.
- Korsgaard, Christine. (1996) *The Sources of Normativity*. Cambridge: Cambridge University Press.
- , (2008) *The Constitution of Agency: Essays on Practical Reason and Moral Psychology*. Oxford: Oxford University Press.
- , (2009) *Self-Constitution: Agency, Identity, and Integrity*. Oxford: Oxford University Press.
- Kvanvig, Jonathan. (2003) *The Value of Knowledge and the Pursuit of Understanding*. Cambridge: Cambridge University Press.
- Lavin, Doug. (forthcoming) "Pluralism about Agency."
- Manne, Kate. (forthcoming) "Democratizing Humeanism." *Oxford Studies in Metaethics*.
- Markovits, Julia. (2013) *Moral Reason*. Oxford: Oxford University Press.
- McGrath, Sarah. (2010) "Moral Expertise as a Puzzle for Moral Realism." *Ethics*.
- Nagel, Thomas. (1986) *The View from Nowhere*. New York: Oxford University Press.
- Parfit, Derek. (2011) *On What Matters*. Oxford: Oxford University Press.
- Peacocke, Christopher. (1992) *A Study of Concepts*. Cambridge: MIT Press.
- Pritchard, Duncan. (2008) "Knowing the Answer, Understanding and Epistemic Value." *Grazer Philosophische Studien* 77 (1):325-339.
- Scanlon, T.M. (2013) *Being Realistic about Reasons*. Oxford: Oxford University Press.
- Schafer, Karl (2013) "Perception and the Rational Force of Desire." *Journal of Philosophy* 110:5 (2013): 258-281.
- . (2015a) "Realism and Constructivism in Kantian Metaethics 1: Realism and Constructivism in a Kantian Context." *Philosophy Compass*.
- . (2015b) "Realism and Constructivism in Kantian Metaethics 2: The Kantian Conception of Rationality and Rationalist Constructivism." *Philosophy Compass*.
- . (2016) "Practical Cognition and Knowledge of Things-in-Themselves." in *The Idea of Freedom: New Essays on the Interpretation and Significance of Kant's Theory of Freedom*, D. Heide and E. Tiffant (eds.), Oxford: Oxford University Press, 2016.
- . (forthcoming) "Rationality: The Capacity for Understanding"
- Schroeder, Mark. (2009) *Slaves of the Passions*. Oxford: Oxford University Press.
- , (2014) *Explaining the Reasons We Share: Explanation and Expression in Ethics, Volume 1*. Oxford: Oxford University Press.
- Setiya, Kieran. (2003) "Explaining Action." *Philosophical Review*.
- , (2009) *Reasons without Rationalism*. Princeton: Princeton University Press.
- Shah, Nishi. (2003) "How Truth Governs Belief." *Philosophical Review* 112 (4): 447-482.

Shah, Nishi and Velleman, David. (2005) "Doxastic Deliberation." *Philosophical Review* 114 (4): 497-534.

Sliwa, Paulina. (2015) "Understanding and Knowing." *Proceedings of the Aristotelian Society* 115: 57-74.

Smithies, Declan. (2015) "Why Justification Matters." in D. Henderson & J. Greco (eds.) *Epistemic Evaluation: Purposeful Epistemology*. Oxford: Oxford University Press: 224-244.

Svavarsdóttir, Sigrún. (2008) "The Virtue of Practical Rationality." *Philosophy and Phenomenological Research* 77 (1): 1-33.

Smith, Michael. (2012) "Agents and Patients, or: What We Learn About Reasons for Action by Reflecting on Our Choices in Process-of-Thought Cases." *Proceedings of the Aristotelian Society* 112.3: 309–331.

———. (2013) "A Constitutivist Theory of Reasons: Its Promise and Parts." *LEAP: Law, Ethics, and Philosophy* 1:9-30.

Street, Sharon. (2008) "Constructivism about Reasons." *Oxford Studies in Metaethics*, 3. Ed. Russ Shafer-Landau. Oxford: Clarendon Press.

----- "What is Constructivism in Ethics and Metaethics?" *Philosophy Compass* 5(5): 363-384.

Thompson, Michael. (2008) *Life and Action: Elementary Structures of Practice and Practical Thought*. Cambridge: Harvard University Press.

Thomson, J.J. (2008) *Normativity*. New York: Open Court.

Tiffany, Evan. (2012) "Why Be an Agent?" *Australasia Journal of Philosophy* 90.2: 1-11.

Velleman, David. (2000) *The Possibility of Practical Reason*. Oxford: Clarendon Press.

----- (2009) *How We Get Along*. Cambridge: Cambridge University Press.

----- (2013) *Foundations of Moral Relativism*. Open Book Publishers.

Way, Jonathan and McHugh, Conor. (forthcoming) "Fittingness First." *Ethics*.

Wedgwood, Ralph. (2007) *The Nature of Normativity*. Oxford: Oxford University Press.

———. (2015) "The Virtue of Rationality", *Analytic Philosophy*.

Zagzebski, Linda. (2001) "Recovering Understanding." in M. Steup (ed.) *Knowledge, Truth, and Duty: Essays on Epistemic Justification, Responsibility, and Virtue*. Oxford: Oxford University Press.