Teleology and Normativity

Matthew Silverstein

This is the penultimate draft of a paper published in *Oxford Studies in Metaethics* 11 (2016). Please cite the published version.

1. INTRODUCTION

Why should the ladder you're building have sides? Because if it doesn't have sides there will be no place to put the rungs. Why should the ladder you're building have rungs? Because if it doesn't have rungs it won't be able to help anyone climb up and down. Why should the ladder you're building be able to help someone climb up and down? It *just should*—that's what a ladder is *for*. If you don't understand why a ladder should be able to help someone climb up and down, then you don't understand what a ladder is!

Here we appeal to the *function* of ladders to put a stop to a regress of "Why?" questions. We have arrived at a "Why?" question that we can answer without prompting yet another such question—a "Why?" question that is settled by the functional nature of the object you're building.

A similar regress looms whenever we attempt to justify our actions, and some philosophers hope that we can put a stop to that regress by applying a more general version of the strategy we use to halt the regress in the case of your ladder-building. This more general strategy appeals not to the functional nature of some specific object or activity, but rather to the function or aim of *agency* itself. Proponents of this general strategy hope that a philosophical investigation into the teleological structure of agency will uncover norms or standards that authoritatively govern us merely in virtue of the fact that we are agents. As agents, we will be

subject to these norms in much the same way that ladders are subject to the norm of helping people climb up and down.

This strategy is a version of *constitutivism*: it seeks to locate the metaphysical foundations of ethics in facts about what is constitutive of agency. A successful constitutivist theory would provide a principled way to halt the regress of justification. It would allow us to meet challenges to our fundamental normative principles by appealing to a nonnormative account of the nature of agency. It promises to reveal not merely which norms have genuine authority, but also what their authority consists in. Constitutivism has therefore garnered a fair bit of attention in recent ethical and metaethical discussions. Yet despite the surge in interest, the philosophical move at the heart of constitutivism remains surprisingly and disappointingly mysterious, for it is not at all clear how we can extract normative conclusions about our reasons for action, or about what we ought to do, from nonnormative premises about the nature of agency. How, in other words, can we derive an "ought" from an "is"? As Hume famously observed, such derivations seem "altogether inconceivable."[1]

Constitutivists who appeal to the teleological structure of agency claim that *functions* or *aims* can help us traverse the "is"-"ought" gap. But these constitutivists must confront a variation on Hume's challenge: How can we derive normative conclusions from merely teleological premises about the function or aim of agency? These teleological premises cannot themselves be normative. If they were, then the resulting view would not fulfill the promises of constitutivism, since it would not locate the foundations of ethics in *nonnormative* facts about the nature of agency. How, then, can we move from teleology to normativity? How do facts about functions or aims bear on facts about what we ought to do or about what we have reason to do? Constitutivists have not ignored these questions, of course, but I shall argue that the answers they have offered so far are unsatisfactory, or at least incomplete. More specifically, I shall investigate two different ways in which prominent constitutivists have conceptualized the teleological structure of action, and thus two different ways of thinking about the connection between teleology and normativity. Both of these approaches fail, or so I shall argue. The constitutivist project can nevertheless be salvaged, but only if it is supplemented with a reductive metanormative account of normativity or reasons for action.

1   Hume (1978: 302).

2. KORSGAARD'S ARISTOTELIAN APPROACH

In her recent book *Self-Constitution*, Christine Korsgaard develops a constitutivist approach according to which constitutive norms are somewhat like Aristotelian *forms*. Following Aristotle's account of the form and matter of artifacts, Korsgaard contends that membership in a kind is determined by the candidate member's teleological organization. Belonging to a particular kind consists simply in having the appropriate *function*. The sort of teleology Korsgaard intends to invoke here is fairly mild. Objects are not endowed with purposes or functions from the outside—by, say, a creator or designer. Instead, an object's function—and thus the kind to which it belongs—is determined entirely by what Korsgaard calls its "internal teleological organization."[2] In other words, an object's kind is determined by what that object *does*—by its "characteristic activity" or work.[3] So, for instance, something counts as a house if and only if it is internally organized so that it can serve as a habitable shelter. Aristotle called this both the *form* and the *essence* of a house. The idea is that being a house just *consists in* being a habitable shelter.

Korsgaard then claims, again following Aristotle, that an object's teleological organization also determines or underwrites its normative status. Embracing what she calls "the ancient metaphysical thesis of the identification of the real with the good," Korsgaard argues that a thing's constitutive teleological organization "gives rise" to "constitutive standards" or norms— that is, normative standards that govern a thing "simply in virtue of its being the kind of thing that it is."[4] More specifically, a thing's constitutive function or form *just is* its constitutive norm. This transition from talk of functions and forms to talk of norms is licensed, according to Korsgaard, by the way in which a thing's constitutive teleological organization can definitively answer normative questions. Why should I add a roof to the house I'm building? Because then it will be able to serve as a habitable shelter. Why should my planned house be able to serve as a habitable shelter? Because that's simply what houses are. To ask why a house should be a

2   Korsgaard (2009: 38).

3   Korsgaard (2009: 27).

4   Korsgaard (2009: 28).

habitable shelter is to betray a lack of understanding of the nature of houses. Such a question just "does not make sense."[5]

Korsgaard contends that we can draw a similar conclusion with regard to action. "Action," she claims, "is self-constitution."[6] It is activity whereby one constitutes oneself. This is the function of action, and it therefore determines what action consists in—what it is for activity to count as action. Moreover, from the fact that self-constitution is the function or form of action it follows that it is also its constitutive standard or norm. Just as it does not make sense to question why a house should be able to provide shelter, so it does not make sense to ask why an action should constitute its agent. That's just what actions do! Once we understand the nature of action as self-constitution, there is no longer any question as to why one should act so as to constitute oneself. The nature of action settles the matter; no further justification is required— or even possible.

I have no interest in challenging Korsgaard's account of action here. My concern is not with her view about the nature of agency (namely, that action is self-constitution), but rather with her broadly Aristotelian approach to teleology and normativity. The worry this approach faces is a familiar one: if the authority of the standards governing action springs from what is constitutive of agency, then any agent will necessarily act in accordance with those standards merely by virtue of being an agent. There seems to be no room for the possibility of bad or incorrect actions.

There will, of course, be various kinds of badness that Korsgaard's view can accommodate. For instance, there is a familiar sense in which an ugly house is a bad house. We might all agree that someone who builds an ugly house is to that extent a bad house-builder. Here what underwrites the charge of badness is not the constitutive norm of houses, but rather some other norm—an aesthetic norm. There is no mystery about how houses can be bad in this sense, since houses might be subject to all sorts of norms—aesthetic, pragmatic, and even moral. But this is not the kind of badness constitutivists like Korsgaard are interested in. Explaining why something is a bad house by appealing to a norm other than the constitutive norm of houses just invites the problematic regress of "Why?" questions. Why should I refrain

5   Korsgaard (2009: 29).

6   Korsgaard (2009: 35).

from painting my house fluorescent purple? Because then it will be ugly. But why should I care about whether my house is ugly? Good question, and one that constitutivists cannot answer.

The kind of badness or error constitutivists are interested in is badness or error *with respect to the relevant constitutive norm*. They are interested in the charge that this house is bad *qua* house—that it fails to meet the constitutive standard for houses. This is the kind of criticism that can halt the regress of "Why?" questions: Why should this house be a habitable shelter? Because that's just what houses are! Korsgaard reserves the term "defect" for the kind of badness or error that is related to something's constitutive norm.[7] We can restate the worry using that term so that it pinpoints the kind of error at issue: How does Korsgaard's constitutivism leave room for the possibility of *defective* actions or agents?

To see the problem, consider the view upon which Korsgaard's is modeled. Aristotle holds that membership in a category or kind is determined by the candidate member's form or essence (*Metaphysics* VII 17). And for both artifacts and natural substances, a thing's form is determined by its purpose or function (*Physics* II 9). Thus, the capacity to chop is what makes something an axe (*De Anima* II 1, 412b10–27; *Parts of Animals* I 1, 642a10), the capacity to shelter people and property is what makes something a house (*Metaphysics* VIII 2, 1043a15), and the capacity for sight is what makes something an eye (*De Anima* II 1, 412b19). Aristotle also identifies the function of a thing with its constitutive standard or norm: "a so-and-so and a good so-and-so have a function which is the same in kind" (*Nicomachean Ethics* I 7, 1098a9).[8] As Korsgaard puts it, "normative principles are in general principles of the unification of manifolds, multiplicities, or, in Aristotle's wonderful phrase, *mere heaps*, into objects of particular kinds."[9] Thus, the standard according to which we judge whether something is an axe (rather than a mere heap) is also the standard according to which we judge whether something is a *good* axe. Similarly, the capacity for sight is not only what makes something an eye; it is also what makes something a *good* eye. The function of an eye is also its constitutive standard or norm.

This "identification of the real with the good" gives rise to a problem for Aristotle (and hence for Korsgaard): it looks like there is no such thing as a truly defective instance of any

7   Korsgaard (2009: 32).

8   All quotations from Aristotle are taken from Aristotle (1987).

9   Korsgaard (2009): 27.

kind. If normative principles are just principles of unification, as Aristotle and Korsgaard suggest, then an axe that cannot chop lacks the unity that would make it a genuine axe. It is just a mere heap. And so an axe that cannot chop is not a bad or defective axe. Instead, *it is not an axe at all*. Aristotle acknowledges as much. After asserting that sight is the essence of the eye, he concludes that "if this fails it is no longer an eye, except homonymously" (*De Anima* II 1, 412b21). A blind man's eyes are thus eyes *in name only*.

Korsgaard's account of action seems to entail a parallel conclusion. Her Aristotelian conception of action's constitutive standard leaves no room for the possibility of actions that run afoul of this standard. Activity through which an agent does not constitute himself is not *bad* action. Since it fails to satisfy action's constitutive norm, it also fails to satisfy action's constitutive condition. Consequently, *it is not action at all*. It is the practical analogue of a mere heap. After all, action *just is* self-constitution, and so a being who fails to constitute himself as an agent is not acting.

This problem is not specific to the particular constitutive norm Korsgaard identifies (namely, self-constitution). It arises instead from her general conception of constitutive standards as Aristotelian forms or essences. On this conception, "every object and activity is defined by certain standards that are *both constitutive of it and normative for it*."[10] There is no room for defective actions because what determines whether something is an action (its function) is *the very same standard* as what determines whether it is a *good* action (its constitutive norm) The same goes for houses and axes—and for anything else governed by this sort of constitutive standard.[11]

---

10   Korsgaard (2009: 32), emphasis added.

11   Korsgaard's discussion of the categorical imperative, which she regards as a genuinely normative principle of practical reason, only confirms this assessment of her view. She claims that an agent who fails to will his maxim as a universal law would have to engage in "particularistic willing." But particularistic willing is impossible, according to Korsgaard:

> There is no difference between someone who has a particularistic will and someone who has no will at all. Particularistic willing lacks a subject, a person who is the cause of these actions. So particularistic willing isn't willing at all. (2009: 76)

Thus the only alternative to acting in accordance with the categorical imperative is not acting at all. There is no such thing as genuine action that violates the categorical imperative. We cannot act badly.

Korsgaard, of course, denies that her view suffers from this problem. She offers the following recipe to distinguish good actions from bad ones: "What makes actions good or bad is how well they constitute you…. People are more or less successful at constituting their identities, and a good action is one that does this well."[12] This sounds like a promising solution to the problem only because it trades on the ambiguity of the crucial word "well." On one understanding of this word, to constitute oneself well is just to do so successfully. But this could not be the basis of an account of bad or incorrect action, since to constitute oneself unsuccessfully is not to constitute oneself at all (and hence not to act at all). On the second understanding of "well," to constitute oneself well is to do so in accordance with some additional, independent norm or standard of evaluation. This *does* allow for the possibility of bad action, but now it is the additional norm—not the constitutive standard—that is pulling the normative weight. Moreover, since this additional norm is not constitutive of action, it cannot play the foundational role that constitutive norms are supposed to play. It stands in as much need of justification as any other external standard for action.

Korsgaard's conception of constitutive norms precludes any satisfying solution to this problem. Either the standard of good action is also the constitutive standard of action in general (in which case bad action is impossible), or we judge the goodness or badness of actions according to an independent standard (in which case bad action is possible but only because we have appealed to a norm that is not constitutive of action and must therefore itself be justified). If self-constitution is what distinguishes actions from mere activity, it cannot also be what distinguishes good actions from bad ones.

Whatever follows from the kind of teleology that interests Korsgaard, then, it is not the normativity of reasons for action. It is not, in other words, the sort of normativity constitutivism is supposed to help us explain. As Korsgaard herself acknowledges, the foundational standards identified by constitutivism cannot be genuinely normative or authoritative if they cannot be violated.[13] To see why, recall that the motivation behind

---

12   Korsgaard (2009: 25).

13   See, for instance, Korsgaard (1996: 161) and (2008a: 40–1). The centrality of the possibility of error to our understanding of normativity is close to an axiom of metaethics. For a list of other philosophers who accept it, see Lavin (2004: 425 n. 1). Lavin himself questions whether this really is a constraint on theories of normativity.

constitutivism is the desire to find norms that cannot sensibly be questioned—norms that are immune to a certain sort of "Why?" question. The hope is that these norms can ground authoritative criticism of agents who flout them. However, as Peter Railton has observed, constitutivist theories that do not allow for the possibility of error have the "unintended effect of pulling the claws of the very criticism" they are supposed to ground.[14] Anyone who runs afoul of agency's constitutive norm fails to count as an agent and hence is not a proper target of rational criticism. The constitutivist can always dismiss such a person as "not one of us," but—as Railton insists—"this is xenophobia, not criticism."[15] To the extent that Korsgaard's view licenses criticism, then, it also ensures there will be no agents worth criticizing.

3. DEGREES OF AGENCY

Korsgaard maintains that the issue is not so straightforward, since membership in a kind comes in degrees on her view.

> But since action requires agency, it follows that an action that is less successful at constituting its agent is to that extent less of an action. So on this conception, "action" is an idea that admits of degrees. An action chosen in a way that more successfully unifies and integrates its agent is more authentically, more fully, an action, than one that does not.[16]

Here Korsgaard suggests that behavior that only partly unifies and integrates its agent is, at least to that extent, action. And if it is action, we can criticize it as bad action. Similarly, an eye

---

14   Railton (2003: 309).

15   Railton (2003: 311).

16   Korsgaard (2009: 25). It is worth noting that this is clearly *not* Aristotle's view. In *Categories* 5, Aristotle states unequivocally that "substance does not admit of a more and a less" (4a7).

> For example, if this substance is a man, it will not be more a man or less a man either than itself or than another man. For one man is not more a man than another, as one pale thing is more pale than another and one beautiful thing more beautiful than another. Again, a thing is called more, or less, such-and-such than itself; for example, the body that is pale is called more pale now than before, and the one that is hot is called more, or less, hot. Substance, however, is not spoken of thus. For a man is not called more a man now than before, nor is anything else that is a substance. (4a1–7)

that sees poorly might still count as an eye, since it does not lack the capacity for sight entirely. We can therefore judge it to be a bad eye.

At best this is only a partial solution to the problem of the impossibility of error, however, for it remains the case that there can be no thoroughly defective actions. There is nothing in the case of action that corresponds to a false belief—that is, something that is both fully a belief and yet altogether incorrect. Korsgaard's view entails that when I constitute myself as an agent sufficiently for my behavior to count as action, my behavior is also and to that extent good or correct.

Actually, I believe that even this partial solution is unsatisfying. A consequence of Korsgaard's approach is that insofar as activity fails to meet action's constitutive norm, it falls outside of the category over which that norm has jurisdiction. In Korsgaard's words, "an action that is less successful at constituting its agent is to that extent less of an action."[17] Similarly, if what makes something a good house is just what makes it a house, then the worse something is as a house, the less of a house it is. But this means that the worse an action (or a house) is, the less it falls under or is subject to *the very standard according to which it is bad*. Put another way, the worse something is, the less weight the charge of badness carries. And so the only way to perform a bad action, on Korsgaard's account, is to act in such a way that action's constitutive norm does not apply to your behavior (since this norm applies only to actions). To the extent that one acts badly, one thereby steers clear of the norm according to which one's behavior is bad.

Korsgaard acknowledges and even embraces this conclusion; she treats it as a feature of her view, rather than a bug. I believe that it diminishes the allure of her constitutivist theory, though. The attraction of constitutivism lies largely in its promise of grounding an authoritative sort of criticism. When we judge that someone acts for a bad reason or that someone performs an unjustified action, we want our judgment to rest on more than some arbitrary norm that itself stands in need of justification. We want our criticism to rest on norms whose authority is beyond dispute—norms that cannot be shrugged off on the basis of ignorance or indifference. Korsgaard's Aristotelian approach grounds this sort of criticism but only in a way that guarantees there will never be any agents worth criticizing. On her view, the more irrationally someone acts, the less that very charge of irrationality sticks. The more worthy of criticism

---

17   Korsgaard (2009: 25).

someone seems, the less weight our criticism carries. This result may be more satisfying than one according to which no authoritative criticism at all is possible, but we should not be content to settle for so little.

4. MORE TELEOLOGY

Korsgaard's Aristotelian approach founders because it leaves no space between action's constitutive function and its constitutive norm. If what makes something an action is also what makes something a *good* action, then there can be no genuinely bad or incorrect actions. Constitutivists must therefore pry these two ideas apart. To make room for the possibility of error, constitutivists must locate a constitutive norm for action that is distinct from action's constitutive function.

One way to do this is by introducing more teleology. Suppose, for example, that whether something has the particular function of a house is determined by more than its *internal* organization. Suppose, in other words, that the teleological nature of a house is at least to some extent determined by *external* factors, such as the intentions of the designer or builder. We can then characterize the functional nature of a house not by saying that it *can* or *does* provide habitable shelter, but rather by saying that it is *for* providing habitable shelter. Something with this sort of function can fail to fulfill that function: something that is *for* providing habitable shelter can easily fall far short of *being* habitable. That is, we can imagine something that is *supposed* to provide habitable shelter but that utterly fails to do so. If, then, we derive the constitutive norm of houses from this more robust sort of teleology, we leave room for the possibility of a thoroughly defective house.

The introduction of this further teleology raises a number of questions. The most immediate is: What kinds of things have functions in this more robust sense? One obvious advantage of Korsgaard's approach is that it appeals to a relatively mild (and thus uncontroversial) sort of teleology. If an object's function is determined solely by its "characteristic activity" or work, then we need not posit a designer or creator in order to understand how objects can have functions. It is enough if there is something those objects *essentially do*. The more robust teleology we are now envisioning is more controversial, though. Artifacts such as houses and axes are the most straightforward examples of this more robust sort of functional kind. Something counts as a house because its *purpose* is to provide a

habitable shelter, and something acquires that purpose through the intentional activity of its creator. Biological kinds such as eyes and hearts might also be examples of functional kinds in this sense.[18] We are all familiar with the idea that eyes are *for* seeing and that hearts are *for* circulating blood.

Is it plausible that *agency* is a functional kind in this more robust sense? I suspect not. Let us suppose nevertheless that agency does have a function in this strong sense. Would the introduction of this additional teleology help? If what makes something an eye is that it is *for* seeing (rather than that it *can* or *does* see), then we can hold that the constitutive norm for eyes is seeing without ruling out the possibility of truly defective eyes. Not everything that is for seeing actually has the capacity for sight. A blind man's eyes will thus count as genuine eyes, and not merely as eyes in name only. The introduction of this additional teleology thus seems like a step in the right direction.

At times Korsgaard at least gestures in this direction:

> The presence of both a noun and a verb in an English sentence is *constitutive* of its being a sentence, that is, of its expressing a complete thought. Yet those of us whose work includes grading papers have all encountered the verbless string of words that *wants* to be a sentence and fails, and yet is not mere gibberish.[19]

> So on this conception, every object and activity is defined by certain standards that are both constitutive of it and normative for it. These standards are ones that the object or activity must at least *try* to meet, insofar as it is to be that object or activity at all.[20]

It is not immediately clear how to make sense of such passages. After all, teleological kinds cannot really *want* or *try* to do anything, Korsgaard's suggestions to the contrary notwithstanding. Strings of words do not want to be sentences; they do not want anything at all. Nor can houses try to be habitable shelters. Perhaps we should not read Korsgaard so

---

18   Here we must tread carefully, of course, for the role of teleology in biological explanations is far from uncontroversial. See Barandalla and Ridge (2011: 367–76) for a helpful discussion of Korsgaard's view in terms of various possible accounts of the notion of a *function*.

19   Korsgaard (2009: 30), second emphasis added.

20   Korsgaard (2009: 32), emphasis added.

literally, though. As I noted above, artifacts have functions in virtue of the intentions of their creators. Strictly speaking, then, what we have is not a string of words that wants to be a sentence, but rather a string of words *intended by its author to be a sentence*. Similarly, what we have is not a pile of bricks trying to be a habitable shelter, but rather a pile of bricks *designed by its builder to serve as a habitable shelter*.

This certainly seems like a step in the right direction, since now we have a view that can explain how something that is fully a house can nevertheless fall short of being a good house: a structure that fails to be habitable may still count as a house if it was intended to be habitable. It should be clear, though, that this appeal to additional teleology cannot be part of Korsgaard's official view.[21] For one thing, it is inconsistent with the "identification of the real with the good," her central metanormative commitment in *Self-Constitution*. For another, she explicitly states that the teleological claims which interest her are "made at the level of the individual object: they are claims about its *internal* organization."[22] Objects that lack the relevant internal organization therefore fail to count as members of the relevant kind: verbless strings of words are not sentences (regardless of the intentions of their authors), structures that are not habitable shelters are not houses (regardless of the intentions of their builders), and eyes that cannot see are not really eyes (regardless of their evolutionary origin).

Moreover, Korsgaard repeatedly asserts that what is constitutive of agency is not merely *trying* to constitutive oneself as an agent, but rather actually *doing* so. The core of Korsgaard's theory of agency is that agents constitute themselves by acting. The alternative view borders on incoherence. If merely *attempting* to constitute oneself were enough to satisfy agency's constitutive condition, then one could meet that condition while failing in one's attempt at self-constitution. In such cases, one both would and would not be a unified self or agent. On the one hand, one *would* count as a unified agent since one would have fulfilled agency's constitutive function, namely, trying to constitute oneself. On the other hand, one would *not* count as an agent since one would have failed in one's attempt to constitute oneself. It appears, then, that

---

21   Here I am disagreeing with those who interpret Korsgaard as defending the view that action *aims* at self-constitution. See, for instance, Enoch (2006: 177, 179), Papish (2011: 456), Setiya, (2013: 3), and Hills (2007: 376).

22   Korsgaard (2009: 38), emphasis added.

Korsgaard could introduce an additional layer of teleology only by abandoning her view that action is self-constitution.

Yet the fact that this path is closed to Korsgaard does not entail that it is closed to all constitutivists. In the next section I shall consider whether there is room for a broadly Aristotelian constitutivism built around this more robust sort of teleology.

5. FUNCTION AND GOODNESS

The distinctively Aristotelian element of Korsgaard's approach is the idea that the teleological structure of agency gives rise to agency's constitutive norm in much the same way that the functional nature of houses gives rise to the norms governing houses. Our normative judgments about agents and actions, in other words, share a conceptual structure and shape with our judgments about the virtues of other functional kinds such as knives and eyes.

Korsgaard is not the only prominent philosopher to have found this Aristotelian conception of normativity attractive. Philippa Foot, Kieran Setiya, and Judith Jarvis Thomson have also developed (or at least explored) broadly Aristotelian versions of constitutivism.[23] Unlike Korsgaard, these philosophers appeal to the more robust sort of teleology described above. For them, something can *have* a function even though it fails to *fulfill* that function: something can be *for* cutting even though it cannot cut at all. They therefore have no problem accounting for the possibility of defective actions. They do, of course, face the additional challenge of showing that agency (or humanity) actually has a function in this strong sense— something which is far from obvious. But I shall not dispute that possibility here. My concern remains not any underlying view about the nature of agency, but rather the broadly Aristotelian approach to constitutivism.

This approach seems so promising because functions seem by their very nature to be normatively significant. Judgments about goodness and badness seem to follow straightaway from facts about functions. Once one knows what knives are *for* (and thus what knives *are*), one knows that knives which cut effectively are good or excellent knives, and that knives which

---

23   See Foot (1961) and (2001), Setiya (2007) and (2013), and Thomson (2008). (Setiya, of course, ultimately rejects constitutivism. I include him here because his presentation of the attractions of the constitutivist approach is so forceful.) Michael Thompson (2008) defends at least elements of the Aristotelian conception of normativity.

cannot cut are bad or defective knives. Foot goes so far as to suggest that the connection between functional natures and normative conclusions is *conceptual*: "'Good knives cut well' must be held to be some kind of analytic proposition."[24] That seems plausible: If you do not know that a good knife is one that cuts well, do you even know what a knife is?

It is difficult to deny, then, that conclusions which look a whole lot like normative judgments follow straightforwardly from premises about functions. Nevertheless, I doubt that a constitutivist theory built solely upon this Aristotelian foundation is viable, because I doubt that the conclusions which follow so straightforwardly from teleological premises are genuinely normative. Functions are *not* by their very nature normatively significant. Conclusions about goodness do follow from premises about functions, of course. We do say, for instance, that a good knife is one that cuts effectively, that a good house is one that provides habitable shelter, that a good eye is one that sees accurately, and that a good heart is one that pumps blood. In fact, I think Foot is correct that these conclusions follow *analytically*: "good" here simply means "fulfills its function," in which case we are still within the sphere of the teleological. When we judge that something is a defective knife, we are saying no more about it than that it is a knife and that it fails to do what knives are supposed to do. What we have here, then, is just teleology masquerading as normativity.

How can we tell? Many philosophers think of normative standards (and normativity more generally) primarily in terms of reasons.[25] And yet nothing about anyone's reasons for action follows from the fact that good knives cut effectively. When we judge that a particular knife is a bad knife merely because it cannot cut, we are not judging that anyone (or anything) has (or lacks) a reason to do anything. Even if I am a knife-maker by trade, the fact that constructing a knife out of cotton rather than carbon steel would result in a defective knife is not necessarily a reason for me to use steel. Nothing about my reasons follows merely from the fact that cotton makes for a very poor knife. Of course, there might be other considerations that together with this judgment yield a normative conclusion. If, for instance, I happen to need a tool that cuts effectively, then the fact that cotton makes for a very poor knife is a reason for me to choose an

---

24 Foot (1961: 48).

25 See, for instance, Raz (1999: 67), who states that "the normativity of all that is normative consists in the way it is, or provides, or is otherwise related to reasons," and Schroeder (2007: 81), who claims that "to be normative, is to be analyzed in terms of reason."

alternative material. More generally, if my circumstances call for a good knife, then I may very well have reason to use steel rather than cotton. But here it is the circumstances that are pulling the normative weight, and so there is nothing resembling the constitutivist strategy at work. We are not deriving a normative conclusion from nonnormative, teleological premises. Instead we are deriving a conclusion about my reasons for action from facts about what would help me meet my needs or achieve my ends. We are, in other words, applying a normative principle along the lines of the principle of instrumental reason. And if our conclusion depends on an appeal to a normative principle, then our strategy for reaching that conclusion cannot be a constitutivist one.

Even if we reject the suggestion that normativity is all about reasons, we can see that Aristotelian conclusions about goodness fall short of genuine normativity. Normative judgments are supposed to be action-guiding, yet the recognition that a knife made out of cotton would be a poor knife need not guide my actions at all, even if I am in the business of making knives. Similarly, normative judgments are typically thought to involve something like an endorsement or recommendation. They are supposed to be *prescriptive*, rather than descriptive. But I am not recommending or prescribing anything when I acknowledge that a good knife is one that cuts effectively. Of course, if you have a pressing need for something that cuts well, then I may very well recommend a quality knife. But once again it is your need that is pulling the normative weight here. By itself, my judgment that some particular specimen is a good knife is just a description. It amounts to no more than the claim that this specimen is a knife and that it does what knives are supposed to do. The same goes for other Aristotelian conclusions about goodness. To see this more clearly, consider some other evaluative judgments that seem to follow straightaway from premises about functions: Arsenic is a very good poison. Muskets are poor guns. Sociopaths make for excellent assassins. My friend is the worst tennis player I have ever seen. None of these judgments involves any sort of endorsement or prescription. I am certainly not prescribing the use of arsenic as a poison. I am not even recommending that you avoid playing tennis with my friend. After all, he is wonderful company, and so it is great fun to play with him, even though he plays poorly. Like my judgment about the knife, these judgments are only apparently normative. Although they

are formulated using words—"good," "poor," "excellent," "worst"—that often have normative significance, that significance is absent here.[26]

Thomson disagrees. She contends that genuinely normative directives follow from Aristotelian claims about goodness—"The pancreas ought to secrete digestive enzymes.… A toaster ought to toast bread"—and that these directives provide a model for directives about people.[27] These conclusions seem strained to me, as they do to many of Thomson's readers.[28] To the extent that such directives are unproblematic, they do not strike me as normative. But I do not want to quibble about the meaning of the word "normative." I am willing to accept that there is a perfectly legitimate use of that term that is broad enough to include within its scope Aristotelian conclusions about goodness as well as the directives which, on Thomson's view, follow from those conclusions. If that is the sort of normativity that interests Thomson, so be it. It is not, however, the sort of normativity that should be of interest to constitutivists. Constitutivists appeal to the teleological structure of agency in order to ground authoritative *criticisms*. They hope to earn the authority to *condemn* certain agents for acting wrongly or irrationally. But as we have seen, Aristotelian conclusions about defective specimens of functional kinds need not involve any criticism or condemnation: I find no fault with my wife when I judge that she would be a terrible assassin. Regardless of how we categorize such Aristotelian conclusions, then, they cannot provide the basis for a constitutivist account of normativity. They lack the relevant sort of normative force.

If this is correct, then our normative judgments about agents and actions do *not* share a conceptual structure and shape with our judgments about the virtues of other functional kinds such as knives and eyes. In other words, the teleological structure of agency does *not* give rise to agency's constitutive norm in the same way that the functional nature of knives gives rise to conclusions about good and defective knives. On its own, at least, the functional nature of agency does not generate any genuinely normative standards. At best, then, the

---

26   Geach, it seems, would have agreed. In criticizing the prescriptivist views of the "Oxford moralists," he insists that "somebody who did not care two pins about cricket, but fully understood how the game worked (not an impossible supposition), could supply a purely descriptive sense for the phrase 'good batting wicket' regardless of the tastes of cricket fans" (1956: 36).

27   Thomson (2008: 207).

28   See, for instance, Wallace (2011).

Aristotelian version of constitutivism is incomplete. This is not the only constitutivist game in town, however. Other prominent constitutivists model the connection between the teleological structure of agency and agency's constitutive norm not on the relation between functions and excellence, but rather on the familiar relation between aims and reasons. As we shall see, these constitutivists face their own difficulties explaining how genuinely normative conclusions follow from the teleological nature of agency.

6. CONSTITUTIVE AIMS

J. David Velleman and, more recently, Paul Katsafanas have developed constitutivist theories that are based on the idea that agency has a *constitutive aim*. According to Velleman, action constitutively aims at an integrative knowledge and understanding of our own behavior. In other words, what makes a bit of behavior an autonomous action is its being regulated by a motive or drive for self-knowledge. From this premise rooted in the philosophy of action, Velleman argues to the normative conclusion that integrative self-knowledge is also the constitutive *norm* or *standard of correctness* for action.[29]

Like the more robust version of the Aristotelian strategy, Velleman's constitutive aim approach leaves plenty of room for the possibility of bad or incorrect actions. Action is activity driven by the aim to understand what one is doing. This is what constitutes action—what all actions have in common. This condition does not require that we actually *attain* such understanding, however. It is merely guidance by the *aim* that constitutes action (and hence agency). Whereas action consists merely in being regulated by the aim of self-knowledge, *correct* action requires actually *achieving* that aim. On the resulting constitutivist theory, the constitutive *norm* of action is different from the constitutive *condition* from which it is derived. And in the space thus opened between them—that is, in the space between the real and the good—resides the possibility of practical error. Behavior that is governed by the relevant aim but that somehow falls short of attaining it satisfies action's constitutive condition but not its

---

29   Velleman first presented his theory of agency in his (1989), although there his constitutivist account of
       reasons for action exists only in embryonic form. Over the next decade Velleman refined and expanded his
       theory in a series of papers (2000b, 2000c, and 2000d). In recent years Velleman has continued to defend
       and refine his theory of agency and the attendant constitutivist account of reasons. See especially (2004a,
       2004b, and 2009).

constitutive norm. The bad or incorrect actions, then, are ones that aim at self-knowledge but that—for whatever reason—fall short of producing it. They are fully actions, but they are also defective or incorrect.

Katsafanas' view differs from Velleman's in many of the details, but it shares the same basic structure: constitutive aims generate or give rise to constitutive norms. Any view with this structure will leave room for behavior that is governed by the aim (and so counts as action) but that fails to achieve that aim (and so counts as incorrect or defective).[30] The constitutive aim approach is not yet in the clear, though. There remains the question of how we are supposed to derive normative conclusions about agency's constitutive norm from merely teleological premises about its constitutive aim. How can nonnormative facts about an aim close normative questions about what we have reason to do? How can we extract normative content from a nonnormative account of the nature of agency? Once again, constitutivism seems to be stranded on the wrong side of the gap between "is" and "ought."

Velleman and Katsafanas both attempt to bridge this gap by appealing to the normative significance of aims. They contend that constitutive *aims* "provide" or "generate" constitutive norms of correctness in a way that other constitutive features may not.[31] According to Velleman, the constitutive aim of action "determines" an internal criterion of correctness for action, and considerations count as reasons for action in virtue of their relation to this criterion.[32] "What lends reasons their force," Velleman maintains, "is an [aim] without which a person is not an agent at all—is not in the business of acting."[33] Since action constitutively aims at self-knowledge, reasons are "considerations out of which we can fashion a description that would embody a knowledge of what we were doing, if we applied that description to

30   Actually, not just any constitutive aim will do. At one point Velleman claimed that action's constitutive aim is simply *autonomy* (2000d: 193; 2000a: 241). He quickly abandoned this idea, though, when it was pointed out to him that were autonomy the constitutive norm of action, all autonomous actions would count as correct. See Clark (2001). Velleman credited Korsgaard for "publicly daring" him to express this mistaken view. Korsgaard's role is not surprising, since—as we have seen—her view suffers from a similar flaw.

31   Velleman (2004b: 283, 288), Katsafanas (2013: 39, 56).

32   Velleman (2000b: 26). See also (2009: 146).

33   Velleman (2004a: 235). Compare (1989: 204–7).

ourselves in the way that would prompt us to behave accordingly."[34] Put another way, reasons are considerations that provide an agent with an explanatory grasp of his own behavior. Yet as Velleman acknowledges, that action has a constitutive aim is merely a fact "about the mechanisms causing and controlling it—in particular, the mechanisms whose causing and controlling it [are] constitutive of its being action."[35] And on Velleman's view, at least, the mechanism in question is just "a fund of labile psychic energy … that motivates various efforts to get the experienced world to hang together."[36] Thus, the claim that action has a constitutive aim looks like a merely descriptive claim, and we are left with the same question: How can we derive normative conclusions about action's constitutive standard of correctness directly from premises about the motives that regulate and control action? How, in other words, can constitutivists manage to pull a normative rabbit out of a merely teleological or motivational hat?

Let me put the problem another way. According to both Velleman and Katsafanas, action's constitutive aim is what lends reasons their normative force. But if the fact that action has a constitutive aim is merely a fact about the motives that produce and regulate action, then the purported force of reasons for action appears to be nothing more than the motivational force that results from triggering or engaging this motive. However, normative force is supposed to be different from motivational force. We are looking for definitive answers to normative questions about what we have reason to do, and it is hard to see how the teleological stories Velleman and Katsafanas tell—stories about how we, as agents, are constitutively motivated— can help. It seems that for any such story, it would remain an open question whether we have any normative reason to be motivated in this way. And so instead of providing an account of reasons for action, Velleman and Katsafanas appear merely to have changed the subject. Of course each contends that his story about action's constitutive aim is *also* a story about the source of the normativity of reasons for action; each claims that the motive that constitutively regulates action also constitutes the agent's responsiveness to reasons. But now our original

34   Velleman (2000b: 27).

35   Velleman (2000b: 17).

36   Velleman (2007: xx).

question returns: Whence comes the normativity of these reasons? In Velleman's words, "what gives them rational authority as opposed to brute motivational force?"[37]

7. SUCCESS AND CORRECTNESS

At times both Velleman and Katsafanas make the transition from a constitutive aim to a constitutive norm or standard of correctness seem easier than it actually is by talking about *success* instead of *correctness*. They suggest that we can derive criteria of success from constitutive aims, which seems quite reasonable.[38] From the fact that my aim is to throw this crumpled bit of paper into the trash bin it seems to follow straightaway that I will succeed if the paper lands in the bin and fail if it winds up on the floor.

As before, though, this entailment is *too* easy. Criteria of success follow so straightforwardly from aims because "aim" and "success" are—or at least can be—both teleological terms. But on this understanding of the word "success," success conditions follow *analytically* from aims. "Success" is just defined as the achievement of one's aim, in which case the fact that Velleman and Katsafanas can derive constitutive criteria of success from constitutive aims is hardly worthy of much excitement: teleology entails teleology. What constitutivists want to derive is something *normative*, and the introduction of the notion of success does nothing to bridge the gap between "is" and "ought." The fact that my action will successfully accomplish what it is constitutively regulated to accomplish tells us nothing about whether I have a reason to perform that action.

Katsafanas often seems not to recognize this gap. He treats "success" as a normative term. Thus he claims that a crucial feature of constitutive standards of success is that they "readily meet challenges to their authority."[39] But a standard of success does not make any claim to authority. It does not tell us what to do or what we have reason to do. It simply tells us what is required in order for us to achieve some aim. It is, in other words, merely a descriptive claim; it

---

37   Velleman (2004b: 293).

38   See, for example, Velleman (2000b: 26; 2000d: 176–82; 2000a: 230–1, 241; 2000c: 246; and 2004b: 283, 287–8), as well as Katsafanas (2013: 39–41).

39   Katsafanas (2013: 40). Later Katsafanas adds: "We need not invoke external facts in order to legitimate their claim to authority" (2013: 40).

contains no normative content at all. Yet throughout Katsafanas' presentation of constitutivism, talk about criteria of success inevitably gives way to talk about reasons:

> Whenever you have an aim, you have a standard of success. Take our aforementioned chess player. Suppose she has the aim not only of checkmating her opponent, but also of enjoying her game. Then we get two standards of success: we can evaluate a particular move with regard to whether the move brings her closer to checkmate, and whether it makes the game enjoyable. These aims can interact with and modify one another: if move A would promote checkmate yet would be boring, while move B would be fascinating yet somewhat more risky, then the player may have reason to make move B. Thus, the reasons induced by the constitutive aim will be one source of reasons among many others.[40]

What Katsafanas never explains is how anything normative—anything involving reasons—is "induced" by a constitutive aim. How do we get from the descriptive or the teleological to the normative? What do aims—constitutive or otherwise—have to do with reasons?

Velleman, at least, acknowledges the gap here. He recognizes that what constitutivists need is not just a standard of success, but a standard of *correctness*. In his more recent work, Velleman addresses the leap from his descriptive account of agency to his normative theory of action's standard of correctness.

> You may object that showing a particular aim to be constitutive of action is not sufficient to show that it sets a criterion of aptness or correctness for action. Achieving the aim may be gratifying to any agent, falling short of it may be frustrating, but these outcomes need not amount to getting things right or wrong, or doing better or worse. Even if considerations relevant to the aim have motivational force for any and every agent, their motivational force might be simply a push or a pull rather than a normative pro or con.[41]

Velleman's reply is succinct.

---

40   Katsafanas (2013: 39).

41   Velleman (2009: 136).

My answer to this objection is that having an aim already establishes a criterion of success or failure, which in turn yields a criterion of correctness for whatever can promote or hinder success. Having an aim establishes a criterion of correctness in relation to that aim.[42]

Here Velleman acknowledges that there are two steps to take: first the move from an aim to a criterion of success, and then the move from a criterion of success to a criterion of correctness. But he fails to see that the second step is crucially different from the first. That teleology entails teleology is uncontroversial. That teleology entails normativity is anything but.

Then again, perhaps there is a sense in which having an aim *does* straightforwardly entail something normative. It is widely believed that one is rationally required to take the necessary means to one's ends. If this belief is correct, then normative conclusions about the rationality and irrationality of various actions do follow merely from the fact that an agent has adopted a particular end.

This seems to be precisely the sort of link between aims and correctness that both Velleman and Katsafanas have in mind. Velleman attempts to illustrate the way in which action's *constitutive* aim provides a *constitutive* or *necessary* standard of correctness by first describing the way in which an agent's *contingent* aims provide *contingent* standards of correctness for his actions.

We can supply a criterion of correctness for any particular action by directing it at a contingent aim—an aim that is contingent in the sense that the action might or might not be directed at it while still retaining its nature as an action. I can flip a switch with the aim of thereby illuminating the room, and then my action can be correct or incorrect as a means to achieving that aim ….[43]

The idea here is familiar: whenever an agent adopts an end, he thereby has a reason to pursue or promote that end. Put another way, by taking on an end an agent makes it the case that achieving that end is a criterion of correctness for his action. Katsafanas, meanwhile, appeals to

42   Velleman (2009: 136).

43   Velleman (2009: 125).

a similar principle: surely we can agree, he suggests, that "if you have an end, you should strive to fulfill it."[44]

Such claims about the normative significance of aims or ends will undoubtedly seem plausible to philosophers of a broadly Humean bent. They cannot serve as a model for the normative force of action's constitutive aim, though, for they draw on the explicitly *normative* principle that one has a reason (or is at least rationally required) to achieve or promote one's aims or ends. The whole point of constitutivism is to derive action's standard of correctness from *nonnormative* facts about the nature of agency. Action's constitutive norm is supposed to be the foundation of *all* practical normative authority. It is supposed to be the point at which justification comes to an end. And so it cannot rely on some other norm for justification—even a norm as seemingly uncontroversial as the principle that we have reason to achieve our aims. Constitutivism is supposed to provide a general account of how normativity arises. If, therefore, the only way to derive a constitutive norm from nonnormative premises about the nature of action involves presupposing some normative claim—even one as "spare" as the claim that we ought to strive to fulfill our aims—constitutivism is in trouble.[45]

If the foregoing is correct, then the move from constitutive aims to constitutive norms is no less problematic than the Aristotelian move from functions to constitutive norms. Both constitutivist strategies are incomplete. Yet I am not ready to conclude that constitutivists must remain forever stranded on the "is" side of the "is"-"ought" gap. In fact, I believe that constitutivists have failed to find a way across that gap not because none exists, but rather because they have been looking in the wrong place. They have tried to account for the normative significance of agency's teleological nature by appealing to the normative force of functions or aims in general. This, I believe, is a mistake. The constitutive aim of action is indeed normatively significant, but not merely because it is an aim. Rather, it is normatively

---

44   Katsafanas (2013: 41). Of course we do not all agree on this point. See, for instance, Broome (1999: 418–9). Later Katsafanas seems to acknowledge this disagreement: "All *but the most ardent Humeans* accept the claim that if you aim at A, you have reason to take steps toward realizing A" (2013: 60), emphasis added.

45   Katsafanas (2013: 41). Later Katsafanas outlines a constitutivist defense of the Humean directive to take the means to our ends (2013: 43–44). This defense appeals to the supposed fact that one cannot aim to bring about an end without aiming to take the means to that end, and so it too already presupposes the normative significance of aims.

significant because it is also the constitutive aim of *practical reasoning*, and practical reason stands in a special relation to reasons for action.

8. PRACTICAL REASONING AND REASONS FOR ACTION

Any account of the teleological structure of action and agency will also be an account of the teleological structure of *practical reasoning*, at least if we understand practical reasoning as reasoning that normally issues in an action (or intention). To say, as Velleman does, that agency aims at integrative self-understanding is to say that practical reasoning is the process of attempting to identify the course of action that will afford the greatest degree of such understanding. When we alight upon that course action, practical reasoning concludes and we act (or at least form an intention).

Once we recognize that the function or aim of agency is also the function or aim of practical reasoning, the route from nonnormative facts about that function to normative conclusions about our reasons for actions becomes somewhat easier to discern. For the idea of practical reasoning seems to be inextricably intertwined with the idea of a reason for action: surely reasons have something to do with reasoning.

Establishing a connection between reasons and reasoning is not sufficient to complete the constitutivist project, of course. The success of the constitutivist strategy hinges on the nature of that connection. More specifically, in order for constitutivists to be able to extract normative content from a teleological account of the aim of practical reasoning, there must be a sense in which our understanding of practical reasoning is *prior to* or *more fundamental than* our understanding of reasons for action. The nature of practical reasoning must *determine* the nature of reasons for action.

Let me elaborate. It should be clear that constitutivists have always been committed to the idea that we can make sense of agency—and thus practical reasoning—in nonnormative terms. They have, in other words, always been committed to the idea that practical reasoning is distinct from normative or ethical reasoning. For if we could understand practical reasoning only in normative terms, then we would be able to explain correct practical reasoning only by first identifying what our reasons for action are. Yet the point of constitutivism is to provide us with an account of reasons for action that does not depend on any normative principles. So, constitutivists must maintain that practical deliberation is carried out under the guise of

something other than the normative. This is not enough to complete the constitutivist project, however. Establishing that practical reasoning is independent of normative reasoning does not show how we can move from premises about the former to conclusions about the latter. To do that constitutivists must defend a more robust claim: they must argue that our understanding of practical reasoning is *prior* to our understanding of reasons for action, or that the latter *depends* somehow on the former. This stronger claim is the additional ingredient required for the constitutivist strategy to succeed. Only by incorporating this further thought can constitutivists make the leap from nonnormative premises about the teleological structure of agency to normative conclusions about our reasons for action.

But what could justify this further thought? Why should we think that practical reasoning is prior to normative or ethical reasoning? I believe that any adequate answer to these questions will require us to wade into the murky waters of metaethics, and in particular into the metaphysics (or semantics) of normativity. More specifically, I believe that any compelling constitutivist story about how to bridge the gap between "is" and "ought" will be rooted in a *reductive* account of normativity—one according to which reasons for action *just are* considerations that bear on practical reasoning. Only with the support of such a reductive metanormative theory can constitutivists show that the teleological structure of agency provides the metaphysical foundations of normativity.

The precise shape of this reductive metanormative theory may depend on how we think of the teleological structure of agency. In other words, different constitutivist strategies may pair well with different reductive hypotheses, and some of these hypotheses might prove more compelling than others. Consider the version of the Aristotelian approach outlined by Kieran Setiya. His reconstruction of the constitutivist strategy is based on a familiar Aristotelian principle:

> EXCELLENCE: When Fs have a defining function or activity, a good F is one that performs that activity or function well.[46]

Setiya acknowledges that the central constitutivist argument cannot be modeled *exactly* on the Aristotelian move from function to goodness. He concedes that the conclusions about goodness which follow from EXCELLENCE are actually just teleological judgments in disguise:

46   Setiya (2013: 3).

"when the nature of a kind is defined in [functional] terms, and where it has a target of which it can fall short, to be good of that kind is to [fulfill that function] in full."[47] Most such conclusions entail nothing about reasons for action. Reasons—and with them genuine normativity—come into the picture only when we apply this principle to *agents*:

> First, agency is defined by a function or activity, an aim or end to which it is directed, just as such—though it is not a condition of acting for reasons that one realize this aim. It follows by Excellence that to be good as an agent is to achieve the aim by which agency is defined. Practical rationality consists in being good *qua* agent. So to be practically rational is to achieve the aim of agency. Since it is a defect of practical reason to be unmoved by a reason to act, this conclusion will constrain what there is reason for us to do. If you can achieve the aim or end of agency while being indifferent to the fact that *p*, this fact is not a practical reason.[48]

Here Setiya moves from Aristotelian (and hence purely descriptive) claims about being good "as an agent" to a genuinely normative conclusion about "practical reason." However, the argument arrives at this conclusion only by way of the claim that "practical rationality consists in being good *qua* agent." The most natural way to understand this claim is as a claim about property identity: the property of being practically rational *just is* the property of being a good agent.

Such a claim certainly seems plausible, especially if we bear in mind the close connection between the function of agency and the function of practical reasoning. Moreover, this particular property identification may appear to be of the sort that is unlikely to raise many metanormative hackles: one type of goodness is being identified with another. But this appearance is misleading. What we have here—whether Setiya realizes it or not—is a reductive hypothesis. As we have already seen, Aristotelian conclusions about being a good agent follow from a functional account of agency only if those conclusions are understood in teleological rather than normative terms. If instead we understand the notion of being a good agent in genuinely normative terms, we are left once again with the mystery of how we can derive normative conclusions about good agents from nonnormative premises about the function of agency. More importantly, however, the strategy Setiya outlines can fulfill its constitutivist

---

47   Setiya (2013: 4).

48   Setiya (2013: 4).

ambitions only if the property with which practical rationality is being identified is a property we can understand in nonnormative terms. If we can understand the property of being a good agent only in normative terms, then Setiya's strategy will have succeeded merely in grounding the normativity of practical rationality in the more fundamental normativity of goodness. But that is not the constitutivist's ambition. The constitutivist's ambition is to ground practical normativity in something descriptive, namely, the teleological structure of agency. This is what makes the view so attractive—and also so challenging.

We are left, then, with the reductive hypothesis that practical rationality consists in fulfilling the function of agency. Why should we think that anything like this hypothesis is true? What does practical rationality have to do with the agency's function? As I suggested above, the connection becomes clearer—and the hypothesis more plausible—once we recognize that agency's function is also the function of practical reasoning. After all, it seems less controversial to suppose that practical rationality consists in fulfilling the function of practical reasoning.

I believe we can make the connection even clearer—and the reductive hypothesis even more plausible—if, following Velleman and Katsafanas, we conceive of the teleological structure of agency in terms of an aim rather than a function. The constitutive aim of agency sets the agenda for practical reasoning. It determines what sort of question we are asking when we engage in practical deliberation, and it thereby also determines what sorts of considerations bear on that question. It determines, in other words, which considerations will bear on practical reasoning. Now a fairly plausible reductive hypothesis presents itself: reasons for action are considerations that bear on the success of practical reasoning. Put another way, reasons for action are considerations that figure in successful practical reasoning—practical reasoning that achieves its aim. This reductive hypothesis is precisely the sort of metanormative view that can bridge the gap between "is" and "ought," or between teleology and normativity. If reasons for action *just are* considerations that bear on successful practical reasoning, and if agency's constitutive aim tells us what those considerations are, then the normative significance of that aim—and of the teleological nature of agency more generally—becomes clear. Teleology entails normativity after all, albeit only with some metanormative assistance.

9. CONCLUSION

Functions and aims seem by their very nature to be normatively significant. The central challenge for constitutivists is to explain this significance—at least in the case of agency—without invoking any prior normative principles. According to the explanation offered by Korsgaard and other Aristotelians, the teleological structure of agency determines what a good agent is in much the same way that the function of a knife determines what a good knife is. And once we know what a good agent is, we can draw further conclusions about what agents ought to do. According to the explanation favored Velleman and Katsafanas, the constitutive aim of agency gives rise to a standard of correctness for action in much the same way that any aim gives rise to a normative standard.

I have argued that both of these explanations fail. The Aristotelian approach articulated by Korsgaard is in the worst shape, since it fails to leave room for the possibility of bad actions—actions that are fully autonomous and yet also fully defective. Aristotelian views that appeal to a stronger sort of teleology avoid this problem, as do the constitutive aim theories proposed by Velleman and Katsafanas. Yet even these views fail to deliver on the promise of constitutivism. Absent some normative principle linking functions or aims on the one hand and norms or reasons on the other, normative conclusion do not follow merely from the teleological structure of agency. Both approaches fail to bridge the gap between teleology and normativity. In light of these failures, it may seem that the only alternatives to these strategies are either to explain the normative significance of the teleological structure of agency by appealing to some foundational normative principle (such as the principle of instrumental reason) or to leave the connection between teleology and normativity unexplained. Neither alternative is consistent with constitutivists' aspirations, and thus we might be tempted to conclude that constitutivism is doomed.

This conclusion would be premature. We *can* extract genuinely normative conclusions from an account of the teleological structure of agency—so long as that account is paired with a reductive metanormative theory linking reasons for action with the aim or function of practical reasoning. I have proposed such a theory: reasons for action are considerations that figure in successful practical reasoning. A reductive view along these lines is bound to be controversial, of course. It will face a number of familiar objections, foremost among them some version of G.E.

Moore's open question argument.[49] I am fairly sanguine about the prospects of overcoming these objections, but there is not sufficient space to justify that attitude here. Instead I shall simply suggest that the reductive hypothesis in question has at least some *prima facie* plausibility. Surely there is an important connection between reasons for action and practical reasoning. And surely this connection has something to do with the aim or function of practical reasoning. If it does, and if we can understand that aim in nonnormative terms, then the basic materials required for the proposed reduction are in place.

It may seem misleading to sell any form of reductionism about normativity as a version of constitutivism, especially since constitutivists often present their approach as an attempt to bypass traditional metanormative disputes. Korsgaard, for instance, describes her view as a "genuine alternative" to standard metanormative theories—one that "stands behind the debate between cognitivists and noncognitivists."[50] Velleman, meanwhile, has serious reservations about the way metaethics is generally practiced: "analyzing the semantics, metaphysics, and epistemology of moral claims will," he maintains, "reveal very little about the fundamental nature of morality."[51] If, however, the arguments of this paper are sound, then the constitutivist project understood as an alternative to traditional metaethics will never get off the ground. Without a reductive theory linking normativity to the aim or function of practical reasoning, there is no way to extract normative conclusions from a nonnormative account of the teleological structure of agency.[52] In other words, without a reductive view of the sort I have proposed, there is no way for constitutivists to bridge the gap between teleology and normativity.[53]

---

49   See Moore (1993: §13).

50   Korsgaard (2008b: 310).

51   Velleman (2009: 157).

52   In Silverstein (2012), I outline an expressivist version of constitutivism that may be able to bypass (rather than bridge) the "is"-"ought" gap. Such a constitutivist view would not require the sort of reductive proposal I have sketched here. Nor, however, would it be an alternative to traditional metaethics.

53   I am grateful to Stephen Darwall, David Dick, David Enoch, Allan Gibbard, Lisa Gourd, Jessica Moss, David Owens, Peter Railton, Kieran Setiya, Nishi Shah, Joshua Silverstein, Phyllis Silverstein, Sharon Street, David Velleman, and various anonymous referees for helpful discussions of the ideas in this paper and for comments on earlier drafts. Earlier versions of the paper were presented to the philosophy departments at Amherst College, the University of Cincinnati, and the University of Minnesota, Morris.

REFERENCES

Aristotle (1987). *A New Aristotle Reader*, edited by J. L. Ackrill. Princeton: Princeton University Press.

Barandalla, Ana and Michael Ridge (2011). "Function and Self-Constitution: How to Make Something of Yourself without Being All That You Can Be." *Analysis* 71, no. 2: 364–80.

Broome, John (1999). "Normative Requirements." *Ratio* 12, no. 4: 398–419.

Clark, Philip (2001). "Velleman's Autonomism." *Ethics* 111, no. 3: 580–93.

Enoch, David (2006). "Agency, Shmagency: Why Normativity Won't Come from What Is Constitutive of Action." *Philosophical Review* 115, no. 2: 169–98.

Foot, Philippa (1961). "Goodness and Choice." *Proceedings of the Aristotelian Society: Supplementary Volume* 35: 45–60.

——— (2001). *Natural Goodness*. Oxford: Clarendon Press.

Geach, P. T. (1956). "Good and Evil." *Analysis* 17, no. 2: 33–42.

Hills, Alison (2007). "Practical Reason, Value and Action." *Journal of Moral Philosophy* 4, no. 3: 375–92.

Hume, David (1978). *A Treatise of Human Nature*, edited by L. A. Selby-Bigge. 2nd ed. Oxford: Clarendon Press.

Katsafanas, Paul (2013). *Agency and the Foundations of Ethics: Nietzschean Constitutivism*. Oxford: Oxford University Press.

Korsgaard, Christine M. (1996). *The Sources of Normativity*. Cambridge: Cambridge University Press.

——— (2008a). "The Normativity of Instrumental Reason." In *The Constitution of Agency: Essays on Practical Reason and Moral Psychology*. Oxford: Oxford University Press.

——— (2008b). "Realism and Constructivism in Twentieth-Century Moral Philosophy." In *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.

Lavin, Doug (2004). "Practical Reason and the Possibility of Error." *Ethics* 114, no. 3: 424–57.

Moore, G. E. (1993). *Principia Ethica*, edited by Thomas Baldwin. Rev. ed. Cambridge: Cambridge University Press.

Papish, Laura (2011). "The Changing Shape of Korsgaard's Understanding of Constructivism." *Journal of Value Inquiry* 45, no. 4: 451–63.

Railton, Peter (2003). "On the Hypothetical and Non-Hypothetical in Reasoning about Belief and Action." In *Facts, Values, and Norms: Essays toward a Morality of Consequence.* Cambridge: Cambridge University Press.

Raz, Joseph (1999). "Explaining Normativity: On Rationality and the Justification of Reason." In *Engaging Reason: On the Theory of Value and Action.* Oxford: Oxford University Press.

Schroeder, Mark (2007). *Slaves of the Passions.* New York: Oxford University Press.

Setiya, Kieran (2007). *Reasons without Rationalism.* Princeton: Princeton University Press.

———(2013). "Murdoch on the Sovereignty of Good." *Philosophers' Imprint* 3, no. 9: 1–21.

Silverstein, Matthew (2012). "Inescapability and Normativity." Journal of Ethics and Social Philosophy 6, no. 3: 1–26.

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

Thomson, Judith Jarvis (2008). *Normativity.* Chicago: Open Court.

Velleman, J. David (1989). Practical Reflection. Princeton: Princeton University Press.

———(2000a). "Deciding How to Decide." In *The Possibility of Practical Reason.* Oxford: Clarendon Press.

———(2000b). "Introduction." In *The Possibility of Practical Reason.* Oxford: Clarendon Press.

———(2000c). "On the Aim of Belief." In *The Possibility of Practical Reason.* Oxford: Clarendon Press.

———(2000d). "The Possibility of Practical Reason." In *The Possibility of Practical Reason.* Oxford: Clarendon Press.

———(2004a). "Précis of *The Possibility of Practical Reason.*" *Philosophical Studies* 121, no. 3: 225–38.

———(2004b). "Replies to Discussion of *The Possibility of Practical Reason.*" *Philosophical Studies* 121, no. 3: 277–98.

———(2007). "Introduction to David Hume Series Edition." *Practical Reflection.* Stanford: CSLI Publications.

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

Wallace, R. Jay (2011). "'Ought,' Reasons, and Vice: A Comment on Judith Jarvis Thomson's *Normativity.*" *Philosophical Studies* 154, no. 3: 451–63.