

What's in an Aim?

Keshav Singh

Penultimate draft – forthcoming in *Oxford Studies in Metaethics*, vol. 17

Introduction

Metaethical constitutivists seek to ground normativity in facts about what is constitutive of agency. One strand of constitutivism locates the foundations of normativity in *constitutive aims*. This strand of constitutivism holds that actions, as well as perhaps a variety of attitudes, have constitutive aims that ground normative standards for them. Defenders of this strand of constitutivism, such as David Velleman (2000, 2004a, 2004b) and Paul Katsafanas (2011, 2013, 2018), tend to conceive of constitutive aims teleologically. In this teleological sense, for some action or attitude *A* to have a constitutive aim *C* is for *C* to be a goal or end such that it is constitutive of *A* that *A* is regulated by the agent's pursuit of *C*.

In this paper, I show that the teleological conception of constitutive aims is inadequate for the constitutivist project. If the appeal to constitutive aims is to succeed, constitutivists need a different conception thereof. I provide such an alternative by sketching a commitment-based conception of constitutive aims. Unlike the teleological conception, the commitment-based conception yields a unified constitutivism that delivers authoritative normative standards for both actions and attitudes.

The structure of the paper is as follows. In §1, I explain the teleological conception of constitutive aims. In §§2-4, I present three challenges to the teleological conception: (1) that it fails to deliver authoritative normativity, (2) that it is not generalizable to attitudes, and (3) that it conflates the aims of agents with the aims of actions (or attitudes) themselves. In §§5-7, I develop my commitment-based alternative to the teleologism and explain why I find it more promising.

1. The Teleological Conception

What does it mean to say that action has a constitutive aim? For metaethical constitutivists who rely on the notion, it has meant that it is constitutive of action that the agent has a particular aim,

in the sense of a goal, toward which the action is directed. In Velleman's words, a constitutive aim is "an aim with respect to which behavior must be somehow regulated in order to qualify as action" (2004a, p. 234). Throughout his work on the subject, he refers to the candidates for a constitutive aim of action as *goals*. So, he understands constitutive aims *teleologically*. Action is constituted as such partly by the agent's directing her behavior toward a particular goal – on Velleman's view, the goal of having an integrative knowledge of what we are doing. Velleman argues that its being action's constitutive aim makes integrative self-knowledge the constitutive standard of success for action as well.

Katsafanas also understands constitutive aims teleologically. In introducing the notion of a constitutive aim, he uses the example of chess: "part of what it is to play chess is to aim at checkmating your opponent. This aim simply must be present in order for a series of movements to count as an episode of chess-playing" (2013, p. 1). This example makes it clear that, like Velleman, what he has in mind is a goal – a motivational state directed at some object one wants to attain. For an activity to have a constitutive aim in this sense is for it to be constitutive of participating in that activity that one's participation is regulated by such a motivational state. On Katsafanas's view, action aims in this sense at pursuing ends and overcoming resistance in the course of doing so.

While Velleman and Katsafanas differ on what the constitutive aim of action is, their theories have the same general structure. Both identify a constitutive aim, in the teleological sense, and claim that it gives rise to a constitutive standard for action that's authoritatively normative. And both argue that this standard is a standard of *success* – if action constitutively aims at some goal or end, then it is constitutive of action that an action succeeds if and only if it attains that goal or end. Finally, both argue that this edifice can ground authoritatively normative reasons for action, because such reasons are generated by a standard of success to which action is by its nature subject.

I'm highly sympathetic to the constitutivist project, and its promise of grounding normative standards for actions and attitudes in their constitutive features. But I think the constitutivist structure

propounded by Velleman and Katsafanas fails, because it rests on a mistaken conception of constitutive aims. In the following three sections, I'll present three challenges to the teleological conception of constitutive aims. Unless teleologists can develop persuasive responses to these three challenges, constitutivists should abandon the teleological conception.

2. Teleology and Authoritative Normativity

It has become increasingly common in metaethics to distinguish authoritative from merely formal normativity.¹ Some norms, such as moral and evidential norms, are thought to be authoritative, while others, such as norms of chess and etiquette, are thought to be merely formal. Merely formal norms are so called because they have the formal structure of norms (oughts, obligations, etc.), but do not by their nature have authority for agents. It is difficult to say what exactly the authority possessed by authoritative norms and lacked by formal ones amounts to, and it is tempting to gloss authoritative normativity metaphorically by saying that authoritative norms “bind us” in a way that merely formal ones do not. For the purposes of this paper, I will try to make do with such a gloss, along with paradigm examples of authoritative and formal norms.

A non-skeptical metaethical theory should deliver authoritative normativity, not just formal normativity. Defenders of the teleological conception (hereafter *teleologists*) think it accomplishes this. But I think their view does little more than presuppose authoritative normativity. Teleologists contend that constitutive aims deliver authoritative normativity because they generate constitutive standards of success for action. As they understand constitutive aims, for action to have some constitutive aim C is for it to be constitutive of action that the agent has some motivational state oriented toward C that

¹ See McPherson (2011), Baker (2017), Woods (2017), and Wodak (2019).

regulates the action-constituting behavior. This generates a standard of success because, when we attain the object of our motivation, this constitutes a certain form of success.

Authoritative norms are often characterized as having “normative force.” For Velleman, “what lends reasons their force...is an inclination without which a person is not an agent at all - is not in the business of acting” (2004a, p. 235). But why should we think the fact that having certain inclinations is a necessary condition for agency generates authoritatively normative reasons to follow those inclinations? Perhaps the authority of the standard of success generated by action’s constitutive aim is supposed to be explained by the universality of this aim for agents. But this seems to answer the wrong question. The question it answers is, “why does this standard of success apply to all agents, and not just some?” The question at hand in grounding normativity in constitutive aims is, by contrast, “why does this standard of success have authority for anyone at all?”

So, the universality that comes from the standard’s being constitutive of agency does not by itself explain why this standard of success is normative. At some points, teleologists seem to think they have delivered normativity simply because ‘success’ is a normative notion. But this amounts to little more than sleight of hand. As Matthew Silverstein (2016) notes:

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 (p. 231).

To put the point another way: if we understand success teleologically, then to succeed *just is* to attain the object of one’s motivation. But this sense of success seems merely descriptive. That one would succeed in this sense if one attained one’s goal does not yet show that one has reason to act in any particular way. Using the word ‘success’ to describe the attainment of a goal evokes the formal structure of normativity, but it does nothing to deliver authoritative normativity. If there is a sense of

‘success’ that’s authoritatively normative, it doesn’t seem to be the sense at play when teleologists derive standards of success from constitutive aims.

Why, then, are teleologists convinced that they can get something normative out of these motivations that are constitutive of action? As Velleman himself puts the question, “what gives [constitutive aims] rational authority as opposed to brute motivational force?” (2004b, p. 293). One possibility is that teleologists are not trying to get something normative out of something purely descriptive at all. Take Katsafanas, who seems not to be trying to ground normativity in something non-normative, but rather to ground substantive norms in antecedent structural ones:

Moreover, notice that Success is an exceedingly spare claim. It can serve as a kind of Archimedean point in debates about ethics: we disagree about whether we have reason to be compassionate, whether happiness is more important than duty, whether suicide is wrong, and so forth. But we can set aside this disagreement on substantive ends and agree on this entirely procedural or structural conception of rationality: we can agree that if you have an end, you should strive to fulfill it, while disagreeing about what those ends are (2013, p. 41).

Here, it seems the starting point is not merely a descriptive claim about what motivations are constitutive of action. It is also the normative claim that we have reason, or ought, to try to attain our ends. If this is right, then the work done by constitutive aims is not to ground authoritative normativity in what is constitutive of agency, but rather to ground substantive norms in structural ones.²

Now, one may think that there is nothing surprising or unsatisfying about this. We must help ourselves to some kind of principle that bridges the descriptive and normative, otherwise we’re simply stuck on one side of the infamous is-ought gap. To derive substantive norms from merely structural ones by showing certain aims to be constitutive of action is the best we can do, and all we should expect out of constitutivism. But if this is what teleologists are doing, then they are not providing a

² Velleman, too, seems to advert to a normative starting point when he discusses norm-expressivism in his 2004b response to critics (pp. 293-295).

foundation for normativity at all. If the structure of the view is that we get authoritative substantive norms out of structural norms plus the fact that there are certain ends all agents have, then teleologists are presupposing the authority of the relevant structural norms. As such, they are merely presupposing, not explaining, authoritative normativity. This is not to say that it wouldn't be an important achievement to ground substantive norms in structural ones. But it would be an achievement in first-order normative theory, not a metaethical foundation for normativity.

This sleight of hand by teleologists is especially worrisome because it is not as if there is already an obvious story about the authority of structural norms. Despite what Katsafanas says, it's not an "Archimedean point" that goals or ends are normative in the way he claims. In fact, it's a matter of vigorous debate whether or not structural norms like the instrumental principle are authoritatively normative.³ And even among those who answer in the affirmative, many of their answers ground the normativity of structural rationality in something else normative.⁴ So, not only do teleologists presuppose something authoritatively normative, what they presuppose is a particularly fraught candidate for authoritative normativity. This is doubly unsatisfying if what we're looking for is a metaethical theory that provides a foundation for authoritative normativity.

3. Constitutive Aims of Attitudes

The second challenge to the teleological conception of constitutive aims is that it's inapt to be extended to explain normative standards for attitudes. While it is somewhat natural to think of action as constitutively involving the pursuit of certain goals, this thought is much less natural when it comes

³ For more on this, see Raz (2005) and the citations contained in footnote 13 therein.

⁴ For example, Kieseewetter (2017) and Lord (2018) both account for structural rationality in terms of substantive rationality.

to attitudes. However, the idea that there is a unified story to be told about the foundations of normativity, one that applies to both action and attitudes, remains highly attractive. While it's possible for the teleological conception to provide a foundation for normativity in the case of action without doing so in the case of attitudes, it would be surprising if the foundations were dissimilar in this way. Normative building blocks like reasons, values, and oughts seem to be just as much at play in the realm of attitudes as they are in the realm of actions. So, a unified explanation is preferable to a piecemeal grounding of various normative domains. Given how commonly endorsed the project of unifying normativity is, I'll assume that if the teleological conception can't be extended to attitudes, this a serious strike against it. This section will focus on establishing the antecedent of that conditional.

The attitude to which the teleological conception has most commonly been applied is belief. Indeed, the 'aims' part of constitutive aims may be traceable back to Bernard Williams' (1970) contention that belief aims at the truth. And when Velleman develops his constitutivism, he argues at several points that belief has the constitutive aim of truth in just the same way action has the constitutive aim of integrative self-knowledge. However, Velleman has since rejected the teleological conception of the constitutive aim of belief, in light of arguments made by Nishi Shah.⁵ I mention this in part because it is striking that one of the primary defenders of the teleological conception of constitutive aims in the case of action now concedes that it is not the right way to understand the constitutive aim of belief. It is worth explaining why this is so. Moreover, the applicability of the teleological conception to other attitudes has not been given the same attention. So, I'll attempt to show not just why the teleological conception is inapt for capturing the constitutive aim of belief, but why it's inapt for capturing the constitutive aims of other attitudes as well.

⁵ See Shah and Velleman (2005) and Shah (2003).

On the teleological conception, for belief to constitutively aim at the truth is for it to be constitutive of belief that it is regulated by the agent's goal (or similar motivational state) of attaining the truth. For the teleological conception to be even remotely plausible with regard to belief, this teleological regulation must be interpreted quite broadly. This is because, as Shah notes, "most of our beliefs aren't formed through our intentional attempts to accept the truth" (2003, p. 460). But this is already to depart significantly from the teleological conception as applied to action, which takes aims to be goals in the sense of intention-like motivational states of the agent. In the case of belief, we instead get something like a regulation of the agent's cognitions by a general truth-aimed disposition. This is still recognizably teleological, but less straightforwardly so.

Against this broadened version of the teleological conception, Shah poses what he calls the *teleologist's dilemma*. This dilemma arises when we try to explain a constitutive feature of belief that Shah calls *transparency*, which is that the question *whether to believe p* gives way to the question *whether p*. That beliefs exhibit transparency is intimately related to the fact that, when reasoning about what to believe, considerations that bear on the truth of the proposition in question play an exclusive role in that reasoning.⁶ Now, here's the dilemma. Teleologists must either take the truth-aimed disposition that regulates belief to be strong, or weak. If they take it to be strong, they can explain transparency, and the exclusive role of truth-related considerations in doxastic deliberation. However, if they take the disposition to be strong enough to explain transparency, this rules out the influence of non-truth-aimed mechanisms. But such mechanisms, such as wishful thinking, clearly *do* influence our beliefs – just not by way of doxastic deliberation.

⁶ Shah's remarks on transparency and the exclusive role of truth-related considerations in doxastic deliberation are closely related to Williams' (1970) remarks on the impossibility of deciding to believe something "irrespective of its truth" (p. 148).

On the other hand, if they take the disposition that regulates belief to be weak, they can account for such mechanisms, but they can't explain transparency. If this disposition is just one among many dispositions jockeying to regulate belief, it's clear how non-truth-aimed mechanisms can influence belief. But now the teleologist has no account of why, when it comes to doxastic deliberation, truth-related considerations play an exclusive role. On this horn of the dilemma, it becomes a mystery why, in reasoning about what to believe, the question *whether to believe p* gives way to the question *whether p*. The source of the dilemma, then, is that by broadening the teleological conception of constitutive aims from an intentional one to a merely causal-dispositional one, the teleologist puts himself in an impossible position, because there is a gap between those mechanisms that merely causally regulate belief, and those that rationally or normatively regulate belief.

What's more, this broadening of the teleological conception *worsens* the problem of authoritative normativity. In the case of action, when confronted with the alleged normative insignificance of aims in the sense of intention-like motivational states, the response that such states regulate not just causally, but rationally or normatively, is at least somewhat plausible. This is what gives the project of deriving substantive norms from structural ones some hope. However, if what's involved when it comes to belief is not intention-like motivational states, but regulating dispositions, such a response is not available. It is, I think, even less promising an idea that authoritative truth-norms for belief can be underwritten solely by the fact that belief is causally regulated by truth-aimed dispositions.

To my knowledge, the challenge involving transparency and the teleologist's dilemma has not been extended to the various non-doxastic attitudes that are similarly subject to normative standards. But the challenge is extensible. Consider fear, for example. Take a plausible candidate for the constitutive aim of fear: danger. Just like a belief is correct if and only if the proposition believed is true, it's plausible that a fear is correct if and only if the object feared is dangerous (on some suitable

interpretation of danger). Now, it can't be that what underwrites this constitutive norm is that it is constitutive of fear that it's regulated by the agent's goal of, say, avoiding danger. Just like most of our beliefs aren't formed through intentional attempts to accept the truth, most of our fears aren't formed through intentional attempts to avoid danger. So, if fear constitutively aims at danger, the teleological conception must similarly be broadened.

It's relatively straightforward how the teleologist's dilemma becomes applicable at this point. Assume fear constitutively aims at danger in the sense that it is constitutively regulated by a danger-aimed disposition. Just as deliberation from non-truth-related considerations can't conclude in belief, deliberation from non-danger-related considerations can't conclude in fear (again, on some suitable interpretation of danger). Though it may be controversial, I'll go a little bit further here and assume that, since this analogy holds, fear, like belief, exhibits transparency to its correctness condition. Now, it seems undeniable that non-danger-aimed mechanisms can exert a merely causal influence on our fears. We now have a gap between what can merely causally regulate our fears, and what can rationally or normatively regulate them. Thus, the teleologist's dilemma is off to the races. The teleologist can't account for the distinctively normative role danger plays in regulating fear, while also accounting for the merely causal influence of danger-irrelevant features.

Though it would help my case, I can't go through the full range of non-doxastic attitudes and show that the same holds of all of them. Fear will have to suffice as a test case for now. My hypothesis is that, for any attitude that's analogous to belief in having a constitutive aim that gives rise to a constitutive correctness condition, it will also exhibit transparency to its correctness condition, an exclusive role for considerations that bear on that condition, and a gap between normative regulation and merely causal influence. As such, it will be subject to the teleologist's dilemma. The teleologist can attempt to produce counterexamples to this hypothesis. But in absence of such counterexamples, I

conclude that the teleologist's dilemma applies to all attitudes that have constitutive aims. As such, the teleological conception is inapt to be extended from the case of action to the case of attitudes.

4. Aims and the Location of Normativity

The third challenge to the teleological conception of constitutive aims is that it conflates the aims of agents with the aims of actions and attitudes. In discussions of constitutive aims, it is common to use the construction "A constitutively aims at C," where A is action, or some attitude, and C is whatever it constitutively aims at – for example, "action constitutively aims at self-knowledge," or "belief constitutively aims at truth." This gives the impression that it is the action itself, or the belief itself, that has an aim. But this makes little sense on the teleological conception. If aims are intention-like motivational states, then actions and attitudes can't themselves have aims. Mental states are borne by agents, not by actions, or by other mental states. Indeed, on the teleological conception, it is ultimately the aims *of the agent* that are constitutive of action, belief, and so on. It is not that action and belief themselves aim at anything, but rather that *in* acting or *in* believing, *we* aim at something.

One might think that "A constitutively aims at C" is nothing worse than an innocuous bit of metaphor on the teleological conception. But I think it displays yet another way in which the teleological conception can't do the work we need from constitutive aims. This is because the fact that constitutive aims are aims of the agent, not of action or attitudes themselves, is in tension with a key part of what teleologists themselves recognize as an important part of the constitutivist edifice:

So what's special about constitutive aims? The reasons derived from the constitutive aim differ from these other reasons in that they are *intrinsic* [emphasis added] to the activity in question... Thus, the interesting feature of constitutive aims is that, being inescapable, they generate intrinsic standards of success. (Katsafanas 2013, pp. 39-40).

I agree with Katsafanas about the significance of constitutive aims being intrinsic to action and attitudes. Part of the promise of the appeal to constitutive aims of action and attitudes seems to be

that we can find the sources of normativity for action and attitudes *in the things themselves*.⁷ But on the teleological conception, the sources of normativity for action and attitudes don't seem to be located in the things themselves. Recall that, on the teleological conception, for action to have some constitutive aim C is for it to be constitutively regulated by the agent's goal of attaining C. Since it is a motivational state that causally regulates the action, the agent's goal is distinct and separable from the action itself. So, the property of being causally regulated by the agent's goal of attaining C looks like an *extrinsic* property of action. If this is right, then Katsafanas's claim that constitutive aims are intrinsic to action actually *undermines* the teleological conception.

When it comes to attitudes, the teleological conception similarly renders constitutive aims extrinsic. In the case of attitudes, constitutive aims must be something like dispositions that causally regulate attitudes for certain properties. But if constitutive aims are causes of attitudes, then we again have what looks like a paradigm case of extrinsic properties. It may well be a necessary condition on, say, belief, that it is regulated by the agent's truth-aimed dispositions. But to identify the aim of belief with such dispositions makes truth the aim of the agent (or more precisely, her cognitive systems) rather than the aim of belief itself.

The teleologist might respond that this ignores the fact that the goal or disposition doing the regulating is *constitutive* of action. If it is constitutive of action, then it partly constitutes action, so it is part of the thing itself. This implies that every constitutive feature of something is also an intrinsic property. But this doesn't seem right – surely there can be cases where constitutive features of things are extrinsic properties. For example, it's possible that certain relationships of mine are partly

⁷ I take this to be an intuitive gloss of intrinsicity talk as it's standardly deployed. For example, for something to be intrinsically valuable is for its value to come from something internal to it; for it to be extrinsically valuable is for its value to come from something else altogether. This conception of intrinsic value is famously both attributed to Moore and defended by Korsgaard (1983).

constitutive of who I am. But this would not make these relationships intrinsic properties of mine; relational properties are paradigmatically extrinsic. Moreover, this more robust understanding of ‘constitutive’ doesn’t even seem to be what teleologists are talking about. When they describe what it is for action to constitutively aim at C, they almost always put it in terms like ‘one cannot act without aiming at C.’ The constitutive aim is characterized as a necessary condition for action, not a proper part thereof.

Perhaps I’m putting too much stock in the language of intrinsicness, though. Perhaps all that teleologists mean is that constitutive aims of action are necessary conditions for action, and are therefore inescapable. If inescapability is all they need, then they certainly have some form of that. But this gets back to one of the problems raised in §2. If the argument here is simply that constitutive aims are special because they are non-contingent, and so inescapable, then it fails for reasons already discussed. If the intrinsicness of constitutive aims is part of their normative significance (and I think it is), then it must amount to more than inescapability. This is why it matters whether aims belong to agents, or to actions and attitudes themselves.

Here’s another way of putting the point, that avoids intrinsicness-talk. Slogans of the form “A constitutively aims at C” evoke a certain kind of promise: a promise of locating the normative standards that govern A in A itself. Part of the attraction of adverting to constitutive aims is that it can deliver a view on which the loci of normativity are actions and attitudes themselves. But for this to be the case, there must be some real sense in which the aims in question are aims of action and the various attitudes themselves, rather than simply being aims of agents. If we locate normativity in the aims of agents, then our constitutivism starts to look just like a kind of subjectivism with the added stipulation that certain motivational states are universal among agents. Constitutivism, at least for many of us, was supposed to be much more than this.

This relates to an earlier point about authoritative normativity. Recall that the teleological conception ends up assuming a structural norm connecting our goals with what we have reason, or ought, to do. As such, it doesn't so much provide a foundation for normativity as it does attempt to derive substantive norms from this presupposed structural norm. Part of why the teleological conception ends up this way is precisely that it has to advert to the aims of the *agent*, rather than aims of action or attitudes themselves. But by conflating aims of agents with aims of action or attitudes, teleologists give the false impression that they have managed to locate authoritative normativity in actions and attitudes themselves, when they have done no such thing.

5. The Commitment-based Conception

In the previous three sections, I developed three challenges for the teleological conception of constitutive aims. Unless teleologists can provide persuasive responses to these challenges, the teleological conception turns out to be inadequate for the constitutivist project. In the course of developing these challenges, I hope to have also drawn out some of the explanatory potential of constitutive aims, a potential not realized by the teleological conception. In the remainder of this paper, I want to explore whether that potential can be realized by an alternative conception of constitutive aims. The alternative I'll develop is a commitment-based conception of constitutive aims. My hope is that, unlike the teleological conception, this commitment-based alternative can yield a unified constitutivism that delivers authoritative normative standards for (and located in) both actions and attitudes.

Whereas the teleological conception starts from action, for which a teleological treatment is more natural, the commitment-based conception starts from attitudes – that is, it accounts first for what it is for an attitude to have a constitutive aim. This is a significant methodological difference from the standard approach to constitutivism in metaethics, which takes action as its primary object.

But elsewhere, in epistemology, there is precedent for thinking about the aim of belief on its own terms instead of modeling it after action. As many epistemologists understand the aim of belief, aim-talk relates to the teleological sense of the word only metaphorically. To say that belief constitutively aims at the truth is not to invoke any motivational states, but rather to say that it is constitutive of belief that a belief is correct if and only if the proposition believed is true.⁸ As Ralph Wedgwood puts it, “It is often claimed that beliefs aim at the truth...But this claim is obviously not literally true. Beliefs are not little archers armed with little bows and arrows: they do not literally “aim” at anything.” (2002, p. 267). In contrast to teleologism, this view is sometimes called *normativism* about the aim of belief.

Though normativism is an account of constitutive aims, it isn't a constitutivist metaethical theory, because it interprets the claim that belief aims at the truth as itself a normative claim. So, normativism by itself is not much help when it comes to looking for an alternative to the teleological conception. What is helpful, however, and what comes from looking at attitudes first, is the idea that constitutive aims don't literally involve aims in the sense of teleological mental states. Instead, constitutive aims are properties that attitudes aim at, in the sense that the nature of the attitude sets conditions that the attitude must satisfy in order to succeed *qua* the kind of attitude it is. This makes clear that a non-teleological conception of constitutive aims must explain what it is in the nature of an attitude that fixes this property as the relevant condition, and what makes the standard this generates an authoritatively normative one.

This is enough ground-clearing to allow me to give a basic statement of my commitment-based conception of constitutive aims for attitudes, which I'll refer to it as **CBC_{ATTITUDES}**:

CBC_{ATTITUDES}: For some attitude A to have a constitutive aim C is for C to be a property such that A is constituted by a committal representation of its object as having C.

⁸ For example, see Wedgwood (2002) and Engel (2013).

According to $CBC_{ATTITUDES}$, all attitudes that are subject to normative standards are intentional mental states⁹ that constitutively represent their objects as having certain properties. For all such attitudes, their constitutive aims are fixed by the accuracy-conditions of these representations. Moreover, these are not just idle representations of objects as having these properties, but *committal* ones. A committal representation of an object's having a certain property is one that involves a commitment on the part of the representor to the object's having that property. It is the fact that attitudes are constituted by such representations that fixes which properties they constitutively aim at. And it is the fact that such representations involve commitments on the part of the agent that makes the standards they give rise to authoritatively normative.¹⁰

Take belief, for example. Belief is standardly thought to be an intentional mental state whose object is a proposition. For belief to constitutively aim at the truth, then, is for the belief that p to be constituted by a representation of p as true that involves the agent's commitment to the truth of p . That a belief is correct if and only if it is true is a constitutive and authoritative norm of belief precisely because belief is constituted by a commitment to the truth of the proposition in question. Or take fear. Assume again that fear's constitutive aim is danger. Though it's less obvious what fear's object is, it's relatively uncontroversial that fear is an intentional mental state. So, for fear to constitutively aim at danger is for it to be constituted by a representation of its object as dangerous that involves the agent's commitment to the dangerousness of that object.¹¹ The fact that this representation is accurate

⁹ 'Intentional' in the sense of *intentionality*, not the mental state *intention*.

¹⁰ In a similar vein, Neta (2018) argues that the involvement of such commitments is what separates "rationally determinable conditions" such as actions and the various attitudes, from other conditions, such as digestion.

¹¹ I should clarify here that this applies only to fear in the sense of a full-blown rationally evaluable intentional mental state, in the way that belief is. What I'm saying about fear obviously wouldn't apply to a simple physiological fight-or-flight response. This clarification is likely necessary for all affective attitudes.

if and only if the object is dangerous fixes danger as the constitutive correctness condition of fear, and the commitment involved accounts for the authoritative normativity of that standard of correctness.

This is the basic idea behind $CBC_{ATTITUDES}$. But there is a lot to unpack here, not least the notion of commitment I'm relying on, and how it explains authoritative normativity. I will try to unpack as much as I can in the remainder of this paper. However, as a disclaimer, what I say will be somewhat programmatic.

The notion of commitment is invoked in metaethics with varied meaning. As some invoke it, commitment is itself a normative term. According to Wedgwood (2002), for one to be committed to making some move x is for it to be irrational not to make move x . On such a view, to say, for example, that believing p and $p \rightarrow q$ commits one to believing q is to say that given one's beliefs, it would be irrational for one not to believe q . So, for Wedgwood, the fact that one is committed to something is itself a normative fact, which is explained in terms of its being irrational not to do what one is committed to doing. Moreover, the sense of rationality in terms of which it's explained is a *structural* one – to be committed to something is for it to be *incoherent* not to do it, just like it's incoherent not to take the means to one's ends. This sense of commitment won't do at all for our purposes, for the same reason the teleologists' reliance on antecedent norms of structural rationality won't do. The most one can accomplish with it is to explain some substantive norms in terms of structural ones.

Perhaps more congenial to the view I'm developing is Pamela Hieronymi's notion of a commitment-constituted attitude:

Notice that, whenever one has an attitude that can be formed or revised simply by settling for oneself a question or set of questions (regardless of how the attitude was in fact formed), one is committed to an answer to the relevant question(s). One is committed in the sense that, if one has the attitude, one is answerable to certain questions and criticisms—namely, those questions or criticisms that would be answered by the considerations that bear on the relevant question(s). So, for example, if I believe p , then I am committed to p as true, that is, I am answerable to questions and criticisms that would be answered by the considerations that bear on whether p . (2005, pp 449-450).

I agree with several things in this passage. One is that attitudes are constituted by commitments such that, for example, if I believe p , I'm committed to the truth of p . Moreover, Hieronymi's point that the commitments that constitute attitudes are answers to questions will ultimately be helpful in explaining several things about $CBC_{ATTITUDES}$. And finally, the connection Hieronymi draws between constitutive commitments and what she calls constitutive reasons – “reasons which bear on the question, the answering of which amounts to forming the attitude” – is key for understanding how constitutive commitments give rise to normative reasons for attitudes. However, Hieronymi's gloss of commitment also won't do for my purposes, because, in understanding commitment in terms of answerability, she also makes commitment out to be itself a normative property. Both Hieronymi and Wedgwood use ‘commitment’ in what Shpall (2014) terms the normative sense, as opposed to the volitional sense.

6. Commitment as Conjunction

For constitutive aims to ground normativity, they must *generate* norms, not presuppose them. This may seem like a kind of magic, which is why those writing about constitutivism sometimes compare it to pulling a rabbit out of a hat.¹² This implies that the magic is only apparent, and this must be a kind of sleight of hand, where the normativity was already there all along. But there's at least one theory of commitment that invokes it as a bona fide foundation of normativity: Ruth Chang's voluntarism. As Chang understands it, voluntarism is a view about the source of normativity that's distinct from either externalism or internalism: “According to normative voluntarists, the source of

¹² For example, see Wiland (2012, p. 141) and Silverstein (2016, p. 19). For related discussion, see also Smith (2015).

practical normativity is to be found neither in irreducibly normative facts nor in non-cognitive states towards which we are passive. Normativity is rather borne of activity of the will” (2009, p. 245).

In developing the view she calls *hybrid voluntarism*, Chang (2013) invokes commitment in the volitional sense, not the normative sense. As she puts it, “commitments are exercises of our *normative powers*, the power to confer reason-giving force on something through an act of will” (p. 75, emphasis hers).¹³ And “[w]illing something to be a reason is the activity of placing your will – your very agency – behind its being a reason” (p. 93). Chang likens this activity to stipulating the meaning of the word, another exercise of a power to make something the case just by willing it be so. We might also liken it to promising, which is also sometimes understood as the exercise of a generative normative power.

Commitments, as Chang sees them, are a source of normativity in that they generate something normative without presupposing anything else normative. Something like this is what constitutivists need. But we will need to do considerably more work here if we are trying to ground all of normativity in commitments. Chang’s more restricted ambition is to ground a certain subset of reasons, which she calls voluntarist reasons, in commitments. Among these are the reasons involved in personal relationships and projects. On her model, when we commit ourselves to others, we *will* that their interests be reasons for us to do things. We’ll return to the case of action later; first, we need to understand how this model of commitment can be applied to $CBC_{ATTITUDES}$.

I’ve claimed that it’s constitutive of having some attitude A that one commits oneself to A’s object having a certain property, C, which is fixed by the nature of that attitude. For example, it’s constitutive of believing p that one commits oneself to the truth of p . Now, it works surprisingly well to model commitment to the truth closely on what Chang says about commitments to others. On this model, when one commits oneself to the truth of p , one wills that one’s acceptance of p be correct if

¹³ For more on normative powers, see Chang (2020).

and only if it is true.¹⁴ Moreover, in doing so, one wills that what is in the interest of truth – in other words, considerations that bear on the truth – be reasons to accept p . The former generates the constitutive standard of correctness for belief, while the latter generates truth-related considerations as (constitutive) normative reasons for belief. Because it's constitutive of belief that one makes this commitment, anytime one is in the business of believing, one creates these norms and corresponding reasons. It is in this way that normative standards for belief come out of the constitutive aim of belief.

Before moving on, one more thing needs to be said about the nature of commitments. Certain commitments are by their nature exclusive, in the sense that having them rules out certain other, incompatible commitments. For example, the commitment to another person that's constitutive of a monogamous partnership is exclusive; it rules out making the same commitment to anyone else. The commitments that are constitutive of attitudes like belief are similarly exclusive. For example, the commitment to accepting truths constitutive of belief is incompatible with a commitment to accepting what maximizes happiness. Belief involves, so to speak, a monogamous relationship with the truth. This is a crucial point, because it explains why one can't will into existence further, non-truth-related normative standards for belief.

Let's sum up what we have so far. For some attitude A to have a constitutive aim C is for C to be a property such that A is constituted by a committal representation of its object as having C . To commit oneself to A 's object having C is to will that one's attitude be correct if and only if its object has C , and moreover to will that considerations that bear on whether the object has C be reasons for A . The willing that constitutes A generates both the fundamental norm that A is correct if and only if it has C , and C -related considerations as reasons for A . This is more like actual magic than sleight of

¹⁴ This is in some ways similar to the norm-expressivism developed by Shah and Velleman (2005). While I don't have the space to discuss that view here, their view differs importantly from mine in that they think what's involved is a higher-order acceptance of the correctness norm for belief. For several reasons, I don't think that works.

hand. It is, in a literal sense, a form of *conjunction*.¹⁵ But of course, that's just what the exercise of a normative power is.¹⁶

This does raise a potential issue, though. If we conjure, say, reasons for believing p in making a commitment that's constitutive of believing p , does this imply that we have no reason to believe p until we already believe it? To give a negative answer to this question, as we surely must, I'll revisit Hieronymi on the formation of attitudes as the settling of questions. On Hieronymi's view, for any commitment-constituted attitude, there is some question the settling of which amounts to forming that attitude. For example, to form the belief that p is to settle in the affirmative the question *whether* p . While Hieronymi doesn't say a lot about what questions are implicated in other commitment-constituted attitudes, like fear or regret, a natural extension suggests itself: the relevant question to some commitment-constituted attitude A is the question whether A 's object has the property C specified by its constitutive aim.

If forming an attitude amounts to settling a question, it's plausible that one's commitments begin not when one settles the question, but when one is confronted with it. To genuinely confront the question *whether* p , one must already be committed to answering in the affirmative if and only if p is true. Moreover, one must already be committed to bringing to bear truth-related considerations on one's answer. The commitments that are constitutive of settling the question *whether* p are also constitutive of confronting the question *whether* p . In other words, these commitments are constitutive of doxastic deliberation. Again, this can be extended to other attitudes. Assume again that fear exhibits something like transparency. The question whether to fear something, then, gives way to the question

¹⁵ My use of the language of conjunction here was inspired by Ram Neta.

¹⁶ Moreover, I take the positing of a normative power to be justified in just the same circumstances as the positing of any other power: when positing it seems to be necessary to explain something.

whether it is dangerous. If there is an analogue of doxastic deliberation for fear, which can conclude in one's fearing something, then it must involve confronting the question whether that object is dangerous. This involves commitments to answer the question in the affirmative if and only if the object is dangerous, and to bring to bear danger-related considerations on one's answer.¹⁷

This begins to explain why attitudes are subject to normative standards prior to the formation of those attitudes: the commitments that ground them are constitutive not just of forming the attitude, but of the deliberation or reasoning that concludes in forming the attitude. In other words, for forming some attitude *A* to even be an *option* for the agent, she must will into existence both the standard of correctness for that attitude and the reasons that bear on whether it is correct. This aspect of $CBC_{ATTITUDES}$ will be crucial for understanding how it accounts for transparency and avoids a version of the teleologist's dilemma.

It might be further objected, however, that $CBC_{ATTITUDES}$ still undergenerates reasons, because it entails that if we haven't yet been confronted with a question for which some attitude amounts to an answer, we have no reasons for that attitude. I can only gesture at two responses to this objection. First, it seems undeniable that we can be confronted with questions without realizing it, or without wanting to consider them. When questions are foisted upon us by the situations we are in, we can't simply decide not to entertain them; our initial commitments in such cases are forced. Second, the source of the objection is in some case actually an advantage of $CBC_{ATTITUDES}$. Epistemologists, especially those who think belief aims exclusively at truth, often struggle to explain why we seem to

¹⁷ To be clear, I'm not claiming that affirmatively answering the question "Is X dangerous?" is *sufficient* for fearing X. My claim is rather that answering the question is a *necessary* condition for one's fear to be the conclusion of deliberation or reasoning, because to deliberate or reason about whether to fear something, one must confront the question of whether it is dangerous. Moreover, as mentioned earlier, this applies only to fear in the sense of a full-blown rationally evaluable intentional mental state.

have no reason to believe wholly inconsequential truths.¹⁸ This problem can be extended to other, non-doxastic attitudes as well. $CBC_{ATTITUDES}$ opens up the possibility of a novel explanation of this datum, which is that we have no reasons for such attitudes because we are not confronted with the relevant questions, so we aren't committed to settling them in any particular way.

7. Challenges Revisited

Even with all of the above clarifications, there is still much more to say about how our commitments ground normativity than can be said in this paper. However, I hope to have said enough to provide a basis for a novel and powerful form of constitutivism. On this form of constitutivism, constitutive aims are not any sort of motivational state. Nor are they dispositions that play the same regulatory role as motivational states. Instead, they are the accuracy conditions of our attitudes. Normativity arises not from our motivational states, but from exercises of our normative power to transform those accuracy conditions into correctness conditions. Some might find such normative conjuration an even more difficult pill to swallow than the normative sleight of hand that usually takes place. But I think its explanatory potential is enough to outweigh such worries.

But what of the challenges I raised for the teleological conception of constitutive aims? At the outset of this paper, I promised that the commitment-based conception would make progress over the teleological conception. I'll now try to make good on that promise. Let's remind ourselves what the challenges are. The first challenge is to deliver authoritative normativity. The second challenge is to deliver a unified account of normative standards for both attitudes and action. And the third challenge is to account for how constitutive aims are aims of attitudes and actions themselves, rather than aims of their agents.

¹⁸ See Treanor (2014) and the quotations and references therein.

First, authoritative normativity. There's a difference in structure between the standard story of how motivational states are reason-giving and how I claim commitments are reason-giving. As discussed in §2, motivational states like goals or ends are supposedly reason-giving because of some principle of structural rationality that connects them to what we do (or at least intend). So, such states are reason-giving only insofar as they are the triggering conditions for preexisting normative principles. This is part of why the teleological conception fails to provide a foundation for authoritative normativity. It at best derives substantive norms from structural norms whose authority is presupposed. Commitments, as I understand them, work very differently. As Chang points out, they aren't properly understood as triggering conditions for preexisting normative principles. Instead, they're properly understood as exercises of generative normative powers. If what we are trying to do with a constitutivist theory is explain how normativity arises without presupposing it, then only something like normative powers can do this work.¹⁹

Moreover, it's not an open question whether what's generated by our wills is authoritatively normative, in the way it's an open question whether there's anything authoritatively normative about brute motivational force. This is partly because we can be alienated from our motivational states in a way we can't be from our wills.²⁰ There's nothing essentially agential about brute motivation, but there *is* something essentially agential about willing. Or, in Kantian terms, our wills are autonomous, whereas our inclinations are heteronomous, so it is only through our wills that we can bind ourselves. In a way, the view I've sketched is a thoroughly Kantian constitutivism: it rejects the normative significance of

¹⁹ An interesting point, which I don't have the space to explore here, is that this kind of account does not fit neatly into the distinction between naturalism and non-naturalism, for it neither provides a naturalistic reduction of normative properties, *nor* holds that there is nothing in virtue of which such properties are normative. On this, I depart from Silverstein, who argues that constitutivism must be paired with naturalistic reduction for it to work.

²⁰ Of course, for this to be the case, willing must be more than just another motivational state.

mere inclinations, instead locating normativity in exercises of the will. Moreover, it posits some commitments as inescapable given the kinds of beings we are. But unlike existing forms of Kantian constitutivism, it does not understand these commitments as endorsements of or identifications with normative principles.²¹ Rather, each instance of reasoning or deliberation brings with it its own commitment, and therefore has a normative authority all its own.

Second, a unified account. The challenge for teleologism was that it could not account for attitudes, because it falls prey to the teleologist's dilemma. Because $CBC_{ATTITUDES}$ can account for the difference between what merely causally regulates attitudes and what normatively regulates them, $CBC_{ATTITUDES}$ avoids the dilemma. Attitudes are normatively regulated by the very commitments that constitute them. These constitutive commitments explain why considerations that bear on the correctness of an attitude play an exclusive role in reasoning or deliberation. In other words, they explain why questions about whether to have some attitude A are transparent to questions about whether A's object has the property specified by A's correctness condition. This accounts for transparency. And unlike teleologism, it does so without ruling out the possibility of normatively irrelevant (i.e., merely causal) influences on attitudes). The fact that certain commitments are constitutive of attitudes does not mean we live up to those commitments perfectly.

However, one might worry that, by starting from attitudes, I've created the opposite problem, and my account of constitutive aims can't be extended to actions. My first response to this is that it's much more plausible that actions are constituted by the kinds of representations I've been discussing than it is that attitudes are governed by aims in the teleological sense. In fact, we can recover much of the appeal of teleology by drawing upon a broadly Anscombean theory of action, much as Velleman does. But unlike Velleman, my view is that we must focus on Anscombe's contention that our

²¹ See, for example, Korsgaard (1996).

intentional actions count as such *not* in virtue of being caused by motivational states, but in virtue of being constituted by representations of our behavior as having certain properties.²² When it comes to what the relevant properties are, my Anscombe-inspiration differs less from Velleman's. I think the relevant properties are those that make our actions intelligible (both to ourselves, and others). Because being part of a teleological series that relates means to our ends is a paradigmatic way of being intelligible to us, this recovers the apparently teleological structure of many actions. However, it doesn't entail that having a teleological structure is the only way of being intelligible (nor did Anscombe think it was). Moreover, it doesn't implausibly try to extend this teleological structure to attitudes.

This gives the basic idea of how the commitment-based conception of constitutive aims can be extended to action. Of course, I haven't explained how we get all the right substantive norms for action out of what makes our actions intelligible. In this way, it's much less obvious how to derive normative standards for action than it is for attitudes. But I think this is precisely what we should expect once we get clear on what constitutive aims must be. So, my focus has been on what constitutive aims must be for aim-based constitutivism to work on a structural level, rather than on whether we can get, say, something like commonsense morality out of the constitutive aims of action. This paper is already too ambitious without trying to take on that latter project.

Finally, the third challenge: the location of normativity. On my view, to say that an action or attitude has a constitutive aim is to say that it is constituted by a representation of its object as having a certain property. In this way, the constitutive aim truly is located in the action or attitude itself, because it in a robust sense constitutes the action or attitude. This stands in contrast to the teleological conception, on which the constitutive aim stands in a causal relationship to the action or attitude, and

²² For more on this interpretation of Anscombe, see Singh (2020).

is thus a distinct entity. Of course, the commitments that constitute actions and attitudes are still mental entities – willings. But it's crucial to the commitment-based conception that these willings are part and parcel of the actions and attitudes for which they generate normative standards. This again stands in contrast to the teleological conception, on which the motivational states that are supposed to be source of normative standards are causes of actions and attitudes. The teleologist's story is one on which motivational states give rise to reasons for something else – actions and attitudes. My story is one on which actions and attitudes themselves give rise to the reasons that govern them. In other words, our actions and attitudes give the law to themselves.

Concluding Remarks

I've tried to show that, while the idea of constitutive aims as the foundation of normativity is very attractive, the teleological conception employed by metaethical constitutivists who rely on it is not up to the task. I've sketched an ambitious alternative, the commitment-based conception, that I think can do the job. What I've provided here is just the beginnings of a full theory, and I don't have the space to explore what I think are its many theoretical payoffs. But to mention one example, I think it can provide a solution to the wrong kind of reasons problem, by explaining both why only right-kind reasons are authoritative reasons for attitudes *and* why we can only respond to considerations that, from our perspective, constitute right-kind reasons. Such payoffs will have to be left for future work.

If there's one thing I hope readers will take away from this paper, it's the following: metaethical constitutivists seek to ground normativity in what is constitutive of agency, but the way in which they often do so opts for mere sleight of hand when real magic is needed. We can't provide a foundation for normativity by pulling it from something else that's presupposed to be normative, as the "magician" pulls a rabbit out of a hat. Instead, we need to posit something that genuinely *conjures*

normativity – something like normative powers. That, in my view, is why constitutivists need a commitment-based account like the one I've sketched here.

References

- Baker, D.C. (2017). The Varieties of Normativity. In Tristram McPherson & David Plunkett (eds.), *The Routledge Handbook of Metaethics*. Routledge. pp. 567-581.
- Chang, R. (2009). Voluntarist reasons and the sources of normativity. In D. Sobel & S. Wall (Eds.), *Reasons for Action* (pp. 243–271). Cambridge University Press.
- (2013). Commitments, Reasons, and the Will. In R. Shafer-Landau (Ed.), *Oxford Studies in Metaethics, Volume 8* (pp. 74–113). Oxford University Press.
- (2020). Do We Have Normative Powers? *Aristotelian Society Supplementary Volume*, 94(1), 275–300.
- Engel, P. (2013). In Defence of Normativism about the Aim of Belief. In T. Chan (Ed.), *The Aim of Belief* (pp. 32–63). Oxford University Press.
- Hieronimi, P. (2005). The Wrong Kind of Reason: *Journal of Philosophy*, 102(9), 437–457.
- Katsafanas, P. (2011). Deriving Ethics from Action: A Nietzschean Version of Constitutivism. *Philosophy and Phenomenological Research*, 83(3), 620–660.
- (2013). *Agency and the foundations of ethics: Nietzschean constitutivism* (First edition). Oxford University Press.
- (2018). *Constitutivism about Practical Reasons* (D. Star, Ed.; Vol. 1). Oxford University Press.
- Kiesewetter, B. (2017). *The Normativity of Rationality*. Oxford: Oxford University Press.
- Korsgaard, C. M. (1983). Two Distinctions in Goodness. *Philosophical Review* 92 (2):169-195.
- (1996). *The Sources of Normativity*. Cambridge University Press.
- Lord, E. (2018). *The Importance of Being Rational*. Oxford: Oxford University Press.
- McPherson, T. (2011). Against quietist normative realism. *Philosophical Studies* 154 (2):223-240.
- Neta, R. (2018). Rationally Determinable Conditions. *Philosophical Issues* 28 (1):289-299.
- Raz, J. (2005). The Myth of Instrumental Rationality. *Journal of Ethics and Social Philosophy*, 1(1), 1–28.
- Shah, N. (2003). How Truth Governs Belief. *Philosophical Review*, 112(4), 447–482.
- Shah, N., & Velleman, J. D. (2005). Doxastic Deliberation. *The Philosophical Review*, 114(4), 497–534.

- Shpall, S. (2014). Moral and Rational Commitment. *Philosophy and Phenomenological Research* 88 (1):146-172.
- Silverstein, M. (2016). Teleology and Normativity. In R. Shafer-Landau (Ed.), *Oxford Studies in Metaethics* (pp. 214–240). Oxford University Press.
- Singh, K. (2020). Anscombe on Acting for Reasons. In R. Chang & K. Sylvan (Eds.), *The Routledge Handbook of Practical Reason*.
- Smith, M. (2015). The Magic of Constitutivism. *American Philosophical Quarterly*, 52(2), 187–200.
- Treanor, N. (2014). Trivial Truths and the Aim of Inquiry. *Philosophy and Phenomenological Research* 89 (3):552-559.
- Velleman, J. D. (2000). *The Possibility of Practical Reason*. Oxford University 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.
- Wedgwood, R. (2002). The Aim of Belief. *Noûs*, 36(s16), 267–297.
- Wiland, E. (2012). *Reasons*. Continuum.
- Williams, B. (1970). Deciding to believe. In *Problems of the Self*. Cambridge University Press. pp. 136--51.
- Wodak, D. (2019). Mere Formalities: Normative Fictions and Normative Authority. *Canadian Journal of Philosophy*:1-23.
- Woods, J. (2018). The Authority of Formality. *Oxford Studies in Metaethics* 13.