There are many normative facts: we have reasons to perform certain actions, certain outcomes of actions are better than others, certain people are virtuous whereas others are vicious, and so on. What explains these very different normative facts? As I understand it, constitutivism is an answer to this question. According to constitutivists, normative facts of certain kinds are explained by facts about the constitutive features of something—that is, the features in virtue of which that thing is the kind of thing it is—where the constitutive features might be those of some person, or some action, or some state of affairs, or something else entirely. Constitutivists may disagree about this.

For those who are wary of talk of explanation, constitutivism can be restated in terms of an entailment claim and a claim about the relative fundamentality of features. Constitutivists think that we can derive claims to the effect that people, actions, or states of affairs of a certain kind have some normative feature from claims about something’s constitutive features, and they further think that these constitutive features are more fundamental than the normative features that they entail. Resorting to a creationist metaphor, we can put the point in terms of what was required of God in the act of creation: all God needed to do was to create things with their constitutive features; once he had done that, the relevant normative features came for free.

Constitutivism, so understood, is a broad church. Constitutivism about value, for example, is the view that facts about the value of objects of a certain kind can be explained in terms of facts about the constitutive features of—of what?—perhaps the constitutive features of objects of that kind (Thomson 2008), or perhaps those of valuers (Lewis 1989), or perhaps those of something else entirely, like God (Zagzebski 2004). These very different views all qualify as versions of constitutivism about value, as constitutivism is being understood here, because they all take facts about something’s constitutive features—whether those of objects of a certain kind, or valuers, or God—to be more fundamental than the value that those constitutive features entail.
Though we haven’t yet said anything about what the constitutive features of things might be, it should be clear that, since facts about (say) a valuer’s constitutive features are supposed to be more fundamental than facts about value, it is inconsistent with constitutivism to suppose that those facts themselves require facts about values for their explanation. More generally, the relevant explanatory relations hold only if facts about constitutive features can be explained without alluding to facts about the normative features they explain. This leaves it open, as it should, that the relevant constitutive features can be explained in wholly non-normative terms, but it also leaves it open that they have to be explained, *inter alia*, in terms of normative features, albeit different ones from those they explain.

Moreover, since constitutivism is just one answer to the question ‘What is a normative fact?’, it also leaves it open whether, in those cases in which the constitutive features of something are themselves explained by normative features, those normative features are explained by something else’s constitutive features. As I understand it, constitutivism about certain normative features is thus consistent with anti-constitutivist views about others. Of course, the most ambitious form of constitutivism will explain all normative features in terms of objects’ constitutive features. But constitutivists need not be so ambitious.

A constitutivist who thinks that we can sometimes explain facts about constitutive features in wholly non-normative terms will reject the idea that there is, in full generality, an *is-ought* gap (Pigden 2010). Moreover, such a constitutivist seems right to reject that idea, as there are kinds whose constitutive features entail normative features where the constitutive features of those kinds can be explained in wholly non-normative terms. Think of functional kinds. If the relevant function can be characterized in wholly non-normative terms, then this will entail facts about what properly functioning things of that kind do, which will in turn entail facts about what ideal things of that kind do, what defective things of that kind fail to do, what things of that kind ought to do, and so on. There is no *is-ought* gap in such cases, so there is no *is-ought* gap in full generality.

An example will help fix ideas. Imagine you have aching limbs and cold extremities, so you go to the doctor. He examines you and then wonders aloud, ‘Why isn’t your heart functioning properly?’ or, perhaps, ‘Why isn’t your heart doing what it is supposed to do?’ How are we to explain his use of ‘properly’ and ‘supposed to’? The answer lies in the fact that a heart is something whose function is to pump an adequate supply of blood around the body (an *is*-claim). This entails that a properly functioning heart pumps an adequate supply of blood around the body (an initial *ought*-claim), that a defective heart is one that fails to do so (another *ought*-claim), and that hearts are supposed to pump adequate supplies of blood around the body (yet another *ought*-claim). The doctor’s uses of ‘properly’ and ‘supposed to’ are in this way explained by reference to the heart’s pumping an adequate supply of blood around the body (an *is*-claim).

It should now be clear why I said that constitutivism is an answer to the question ‘What is a normative fact?’ Even those who are convinced that there are irreducibly normative features should agree that, when it comes to facts like the heart’s functioning properly, or being defective, or doing what it is supposed to do, the constitutivist’s explanation is the best available. But if these normative features can be explained in wholly non-normative terms, it cannot be absurd to suppose that all normative features are so explainable, much as the ambitious constitutivist thinks. Constitutivists needn’t be so
ambitious, but since we’re stuck with some constitutivist explanations, parsimony tells in favor of looking for constitutivist explanations wherever we can and, when we find them, making sure that the explanations given are as ambitious as possible.

In the light of this, consider again constitutivism about value. As we saw, David Lewis tells us that for something to be of value is for that thing to be valued by ideal valuers (Lewis 1989). Lewis’s explanation of value initially appears to be normatively laden—what is an ideal valuer?—but we can now see that it may be fully reductive for all that. According to Lewis, valuers are those with the capacity to match their valuings, which he thinks are their intrinsic second-order desirings, with those that would survive full imaginative acquaintance with their objects. Valuers thus value well to the extent that their intrinsic second-order desires are those that would survive the exercise of this capacity, and badly to the extent that they are not, and an ideal valuer is someone who values as well as possible. Lewis’s explanation of value therefore makes no appeal to values, and it explains what an ideal valuer is in functional terms, not constitutivist terms.

We can now characterize a constitutivist view of some normative fact more precisely. What makes the earlier account of facts about defective hearts, and what hearts are supposed to do, a constitutivist account is the fact that the kind heart, being characterizable functionally, generates an ordering of hearts from best to worst according to how well they function. The kind heart is thus what Judith Jarvis Thomson calls a goodness-fixing kind, and the account of defective hearts and what hearts are supposed to do that we give by appealing to the functional characterization of hearts is fully reductive because that ordering, which can be understood as a specification of the functional nature of hearts, can itself be characterized in wholly non-normative terms.

Similarly, what makes Lewis’s account of value a constitutivist account is the fact that the kind valuer, being characterizable in terms of the capacity to have second-order intrinsic desires that match those that would survive full imaginative acquaintance with their objects, also generates an ordering of valuers from best to worst—the kind valuer is a goodness-fixing kind—and it is fully reductive if a specification of that ordering can be given in non-normative terms. Whether this is so will depend on whether we can explain what desires are without presupposing anything normative. But if, for example, to be a desire is to be a psychological state that plays a certain functional role along with beliefs in the psychology of a fully rational agent, and if we could explain what an agent’s being fully rational is without presupposing anything normative, perhaps by supposing that the beliefs and desires of fully rational agents bear certain structural relations to each other, and that psychologies can be ordered from best to worst depending on how similar their structural relations are to these, then the account would be fully reductive. Lewis’s ambitious constitutivism about value would then follow from functionalism about psychology.

In general, an account of some normative fact is constitutivist when the account explains that normative fact in terms of some goodness-fixing kind, and the account is fully reductive when the specification of the ordering that that kind generates can be given in wholly non-normative terms. With this general characterization of constitutivism before us, it should be clear that theorists have offered very different constitutivist accounts of a diverse range of normative facts. Whether their versions are fully reductive is harder to assess, but for many, the aim doesn’t seem to have been reduction, but rather a decrease the number of normative facts that require explanation. Here are some examples.
Consider Thomson’s explanation of facts about what human beings ought to do (Thomson 2008). In her view, among other characteristics, what it is to be human is to have moral capacities like the capacity to be just and generous. The kind human being is therefore a goodness-fixing kind: we can order human beings from better to worse depending on how just and generous they are; we can identify defective human beings as those who lack the virtues of justice and generosity; and we can define what human beings ought to do as those things which are such that, if they knew what differences would be made by their doing those things, they would be defective human beings if they failed to do them. Thomson thinks that what human beings ought to do can in this way be explained in terms of a specification of the goodness-fixing kind human being, a specification which itself makes no appeal to facts about what ought to be done. But since the explanation appeals to facts about agents being just and generous, which are themselves normative facts, her account of what humans ought to do is fully reductive only if we can explain these virtues in non-normative terms.

Or consider Jurgen Habermas’s explanation of what it is for a rule of action or choice to be justified in terms of whether all those who are affected by the rule or choice could accept it in a reasonable discourse (Habermas 1996). This turns out to be a constitutivist explanation of justified rules and choices because it is in the nature of a discourse to be a mode of communication that can more or less reasonable—the kind discourse is therefore a goodness-fixing kind—and because we can define what it is for a rule of action or choice to be justified in terms of what could be accepted in a reasonable discourse. Whether rules of action or choices are justified are in this way supposed to be explained by a specification of an ordering generated by the goodness-fixing kind discourse, a specification that itself makes no appeal to facts about which rules of action or choices are justified. But since Habermas’s explanation appeals to facts about discourses, and discourses are in turn the product of our communicative intentions, his explanation is like Lewis’s in being fully reductive only if we can understand psychology in wholly non-normative terms.

Or, more ambitiously, consider Sharon Street’s claim that what makes claims about what people have reason to do true is the fact that those claims follow from the practical point of view, where the practical point of view is the point of view of someone who values various ends and the means to those ends, and a claim about what they have reason to do follows from this point of view if it is supported by the valuings that would survive if that person’s values were as coherent as possible (Street 2009, 2012). This is a constitutivist explanation because the kind practical point of view generates an ordering of sets of valuings from best to worst depending on how coherent they are—the kind practical point of view is therefore a goodness-fixing kind—and a specification of that ordering makes no appeal to facts about what people have reason to do. Street’s explanation looks to be more reductive, but since valuing is a kind of psychological state, her explanation is like Lewis’s and Habermas’s in being fully reductive only if we can understand psychology in wholly non-normative terms.

It will be helpful at this point if we focus on a specific example. Since we just considered Street’s view, let’s consider other versions of constitutivism about reasons for action, and let’s note how the distinctions made so far help us understand the differences between these different versions, and how attention to these differences, so understood, helps us identify the different versions’ strengths and weaknesses. To anticipate, the crucial questions will be: which goodness-fixing kind is supposed to explain facts about reasons for
action; whether a specification of that goodness-fixing kind presupposes normative facts, and if so, which; and how compelling the explanation provided is. Since the answer to the last question will depend on whether the explanation gets the extension of what we have reason to do right, how compelling any particular version of constitutivism is will depend on the extent to which it survives in reflective equilibrium with our substantive views about what we have reason to do. Given theorists’ disagreements about what we have reason to do, this makes the evaluation of different versions of constitutivism about reasons for action problematic.

The disagreement that will occupy center-stage in what follows is whether moral requirements entail facts about what we have reason to do, and, more specifically, given that on any plausible view of morality we are sometimes morally required to cooperate with others when doing so requires us to give up something that we care about more than cooperation, whether we sometimes have reasons to act contrary to our cares and concerns. Street’s view is striking in this regard because she thinks it follows from her explanation of what we have reason to do that we never have reasons to act contrary to the cares and concerns that would survive as part of a fully coherent set. A perfectly coherent Caligula, she tells us, has no reason at all to do what’s morally required (2009). Is this an objection to Street’s version of constitutivism? That depends on whether you think moral requirements entail reasons for action. At this point, I invite readers to make a mental note of their answer to this question and to score Street’s constitutivism accordingly.

Street’s view is striking for another reason as well. The target of her explanation is familiar facts about what we have reason to do, but she pursues her reductive ambitions by couching her explanations in unfamiliar terms. Though the details are unimportant, both the idea of a practical point of view and the more basic idea of someone’s valuing something turn out to be technical terms, as Street uses them. This adds an extra dimension of difficulty to the evaluation of her view, as we first have to figure out what she takes the practical point of view and valuing to be. We then have to ask ourselves whether she is right that, for all their unfamiliarity when fully spelled out, it turns out that those of us who have reasons for action have a practical point of view and value things, and only then do we get to ask whether the explanation she provides is compelling. It is therefore worth considering some alternative explanations where the choice of goodness-fixing kind is more familiar.

In *Natural Goodness*, Philippa Foot tells us that there is the following strong similarity between living things (Foot 2001). For each species of animal and plant, there is a set of generic claims that spells out what, for that species, is the characteristic way in which its members develop, maintain themselves, and reproduce (see also Thompson 1995). These generic claims spell out what it would be for a member of that species to function properly, and hence which features would be possessed by species members that are functioning properly—these are the relevant plant and animal goods—and what it would be for them to be defective. In the case of animals where the activity necessary for development, self-maintenance, and reproduction cannot be explained in stimulus-response terms, but requires belief-desire explanations, these generic claims will include claims about actions that realize the relevant animal goods, and, in the case of social animals like human beings and wolves, they will also include claims about the cooperative activity needed to realize the relevant animal goods.
With a rich account of distinctively human goods in place, Foot thinks that we can give an account of the human virtues. Human virtues turn out to be those traits of character that a human being needs to have in order to fare well as a human being. At a minimum, this presumably requires that the desires that drive the actions of virtuous human beings have as their contents the human goods—or anyway, contents that stand in some systematic relation to human goods—and, with this account of the human virtues in place, we also have the materials out of which to construct a theory of reasons for action. Reasons for action turn out to be those features that motivate virtuous agents, so once again, these features will turn out either to be, or to stand in some systematic relation to, the human goods that virtuous agents bring into being.

In explaining reasons for action, Foot thinks we must appeal to the fact that human beings are of a kind with other living things, and that things of this kind have a characteristic function, namely, that associated with their development, self-maintenance, and reproduction, a function that can be specified in terms of an ordering from best worst. Human being is thus a goodness-fixing kind. Moreover, Foot thinks that once we specify this ordering we can explain all sorts of normative features: what it is to be a defective human being; what human goods are; what it is for a character trait to be a human virtue; and what the relationship is between human goods, human virtues, and reasons for action. Given that Foot makes no appeal to facts about reasons for action in her specification of this ordering, it follows that her theory meets the explanatory constraint on a constitutivist theory of reasons for action. Whether the specification could itself be explained without appealing to normative notions is an open question—note she too must appeal to beliefs and desires—but giving such a non-normative specification doesn’t seem to be her main concern.

One of the great virtues of Foot’s version of constitutivism, as she sees things, is that it explains why we sometimes have reasons to cooperate even when doing so requires us to act contrary to our cares and concerns. This is because whether or not some action realizes a human good isn’t dependent on whether it realizes something the agent cares about, or would care about if she were coherent. Reasons to cooperate are not hostage to our cares and concerns, but are instead a consequence of the behaviors we have to engage in, as social beings, in order to develop, maintain ourselves, and reproduce. To the extent that these cooperative behaviors correspond to what’s morally required, Foot’s version of constitutivism thus squares with the idea that moral requirements entail reasons for action. This is a striking difference between her constitutivism and Street’s, and it is a direct consequence of her choice of goodness-fixing kind.

Foot’s view is not, however, without its problems. For one thing, the cooperative behaviors Foot tells us we have reasons to engage in correspond only imperfectly to what’s morally required. Imagine that there are extraterrestrials who are as psychologically sophisticated as humans. Foot’s view entails that, just as humans have reasons to realize human goods, such extraterrestrials have reasons to realize extraterrestrial goods. But now imagine a situation in which humans and these extraterrestrials interact, but no course of action available simultaneously realizes both human goods and extraterrestrial goods. If humans and extraterrestrials were each to do what they have reason to do then, given Foot’s explanation of reasons, they would be at loggerheads with each other, each trying to realize their own species-specific good. Foot’s version of constitutivism thus squares with the idea that we have reasons to do what we are morally required to do only
on the assumption that morality doesn’t tell us that there are fair terms on which humans and extraterrestrials could interact in such circumstances, not even fair terms that put bounds on what a competition in the pursuit of their respective species-specific good could consist in. In putting so much emphasis on human goods, Foot’s view thus builds an implausible human chauvinism into morality.

Foot’s human chauvinism has further implications. Though her explanation of reasons allows that we may sometimes have reasons to do something because we care about it, it only allows that this is so when getting what we care about doesn’t come at the cost of human goods. To focus on a striking case, though many people care about having children, it is a commonplace that some don’t, and we ordinarily think that the decision to have children is therefore a matter of personal preference. But now imagine that everyone had the preferences of those who don’t want children. On Foot’s view, independently of whether their having children would help them get other things they care about, or other human goods, everyone would still have reasons to have children. Or imagine people who care very much about the goods of dogs, and cats, and cows, and pigs, and chickens. If they cared so much about these non-human beings’ goods that they sacrificed human goods for their sake, then on Foot’s view they would be doing something they have no reason to do. These too are consequences of Foot’s human chauvinism.

An obvious thought at this point is that Foot chose the wrong goodness-fixing kind in terms of which to explain reasons for action. She should have chosen a kind that didn’t entail a species-specific ranking, and a kind that allows for the possibility that beings of the relevant kind can do well simply by doing things they care about. With this thought in mind, consider Bernard Williams’s explanation of reasons for action (Williams 1980).

Williams explains what agents have a reason to do in terms of what they would be motivated to do if they deliberated correctly. Agents deliberate correctly, according to Williams, only if (i) they have all and only those intrinsic desires they would have if they exercised their imaginations concerning all of the possible objects of intrinsic desires (Williams’s explanation of reasons is, in this respect, like Lewis’s explanation of value); (ii) when their intrinsic desires have indeterminate content, and they could make them more determinate by bringing them into line with other things they intrinsically desire, they do so (to the extent that this sounds like a coherence requirement, Williams’s explanation of reasons is in this respect like Street’s); (iii) they have no false non-normative beliefs relevant to the satisfaction of their intrinsic desires; (iv) they are not ignorant of any non-normative facts relevant to the satisfaction of their intrinsic desires; and (v) they order the actions they take to satisfy their intrinsic desires over time so as to ensure the optimal satisfaction of their (current) intrinsic desires (again, to the extent that this sounds like a coherence requirement, the explanation is like Street’s).

Though Williams doesn’t describe his explanation of reasons for action as a version of constitutivism, it certainly seems to be a version. Williams explains reasons for action in terms of the kind deliberator, where deliberators have the capacity to deliberate, and where deliberators deliberate well to the extent that they exercise this capacity and their psychologies have the features he describes, and badly to the extent that they fail to exercise the capacity and their psychologies lack such features. Deliberator is thus a goodness-fixing kind. Williams’s explanation of reasons is given in terms of a specification of those at the top of the ordering that this goodness-fixing kind generates. Given that the
specification makes no appeal to facts about reasons for action, it follows that Williams's
type meets the explanatory constraint on a constitutivist theory.

In some respects, Williams's explanation is like Street's, but his explanation is unlike
hers in appealing to a familiar kind whose connection with reasons for action is more or
less transparent: the role of deliberation just seems to be the discovery of what there is
reason to do. However, if Williams is right, this is so even though we needn't explicitly
ask ourselves what there is reason to do when we deliberate: look again at (i)–(v). As
with Thomson's and Foot's explanations, whether Williams's explanation is fully reductive
depends on whether we can explain what it is for someone to have desires and beliefs in
wholly non-normative terms. In this connection, it worth remembering the suggestion
made earlier in the discussion of Lewis's view of valuing.

If desire and belief are understood functionally in terms of how they relate to each
other in the psychology of a fully rational agent, then if we could we explain what an
agent's being fully rational is without presupposing anything normative, perhaps by sup-
posing that the beliefs and desires of fully rational agents bear certain structural rela-
tions to each other that generate an ordering of psychologies from best to worst, Lewis's
account of valuing would be fully reductive. What is striking about Williams's (i)–(v) is
that they look like an attempt to describe some of what these structural relations might
be. Though Williams doesn't advertise himself as giving a fully reductive explanation of
reasons for action in constitutivist terms, it does seem to be an attempt to give such an
explanation.

Note that Williams's specification of deliberators is not species-specific—humans and
extraterrestrials have the same deliberative capacities—but instead ranks deliberators
according to how well they meet conditions (i)–(v). Moreover, note that these conditions
allow that we can have reasons to act on our cares and concerns to the extent that these
survive the exercise of our imagination about their objects. Williams therefore has no
problem explaining why people with the relevant intrinsic desires may have no reason
to reproduce, and why they may have reasons to realize non-human goods even when
doing so comes at some cost to human goods. In this respect, Williams's explanation
better captures what we ordinarily take to be the extension of reasons for action. But is
Williams's explanation consistent with the claim that moral requirements entail reasons
for action?

According to Williams, deliberators have the capacity to acquire intrinsic desires
after they fully exercise their imaginations. But since the intrinsic desires they acquire
are explained by facts exogenous to their being deliberators—deliberators have differ-
ent embodiments, upbringings, cultures, and so on, and it is these that explain which
dispositions to acquire intrinsic desires in the light of their imaginings they have—it fol-
lows that even deliberators who deliberate well could end up with very different intrinsic
desires from each other. Williams's explanation is thus like Foot's in making facts about
reasons for action relative, but instead of being relative to species, they are relative to the
individual. The upshot is that, if Williams's explanation is correct, then whether humans
have reasons to cooperate on fair terms with extraterrestrials, and indeed with each other,
depends on whether, after they exercise their imaginations, they have intrinsic desires
that will be satisfied by such cooperation. Williams's and Street's views are in this respect
strikingly similar. Neither allows that moral requirements entail reasons for action. Once
again, I invite the reader to score Williams's explanation in the light of this.
Should those who think that moral requirements entail reasons for action abandon the idea that such reasons can be explained in terms of the constitutive features of deliberators? Not necessarily. When we think of Williams as a constitutivist, his constitutivism consists of two claims: first, that we explain reasons for action in terms of the constitutive features of deliberators, and second, that the nature of deliberation is as specified in (i)–(v). A constitutivist who thinks that moral requirements entail reasons for action could think that the problem with Williams’s explanation of reasons lies not in his choice of goodness-fixing kind, but rather in his specification of the nature of that kind (compare Korsgaard 1986). In particular, perhaps Williams is wrong that the only way we can correct our intrinsic desires when we deliberate is by revising our intrinsic desires in the light of exercises of our imaginations, where this is in turn explained by exogenous factors like our embodiment, upbringing, enculturation, and so on.

Unsurprisingly, Kant can be understood as a constitutivist who thinks that this is so (Kant 1786). Reasons for action are to be explained in terms of a specification of the goodness-fixing kind deliberator, but those who deliberate well additionally exercise the capacity to give themselves the Categorical Imperative, and so come to have motivations that conform to universal laws. What it is to deliberate well, on Kant’s view so understood, just is to be so motivated—in Kant’s terms, it is to have a “good will.” The challenge is then to explain two things: first, what exactly it means to have motivations that conform to universal laws, and second, why the mere fact that a deliberator has such motivations entails that they have reasons to cooperate on fair terms with other deliberators. Moreover, if the explanation is to satisfy the explanatory constraint on a constitutivist theory of reasons for action, these explanations must at no point appeal to facts about what deliberators have reasons to do.

Very roughly, Kant’s answer to the first question is that motivations accord with universal laws when they are motivations to act in ways that all deliberators could act at the same time as a result of their deliberating, and his answer to the second question is that it follows from this that no deliberator could end up with motivations to treat other deliberators in ways that leave them incapable of acting on the motivations that result from their own deliberations. In this way, Kant’s specification of what it is to deliberate is supposed to entail that everyone has a reason to treat deliberators, whether themselves or others, never merely as a means, but also always as an end. If, as seems plausible, this entails having reasons to cooperate with other deliberators on fair terms, then the Kantian explanation of reasons for action is consistent with the claim that moral requirements entail reasons for action. Not only do human deliberators have reasons to cooperate on fair terms with each other, but they also have reasons to cooperate on fair terms with extraterrestrials, and extraterrestrials also have such reasons.

There are, of course, well-known problems with Kant’s answers to each of these questions, and there are well-known Kantian attempts to address these problems (Korsgaard 1986). For present purposes, the important points are, first, that the answers to the two questions do not anywhere appeal to facts about reasons for action, so the explanation meets the constitutivist’s explanatory constraint, and second, that the answers seem to be a further elaboration of the structural relations between the intrinsic desires and beliefs of those who deliberate well. Deliberators have the capacity to move from thoughts about what would be required for motivations to conform to universal laws to corresponding motivations, so a specification of what it is to deliberate well simply amounts

\[\text{\textsc{constitutivism}}\]
to a description of the ways in which thoughts about universal laws and corresponding motivations in fact relate to each other in deliberators who have and fully exercise these capacities. In this respect, Kant’s explanation is like Williams’s. Both are well understood as a contribution to a fully reductive explanation of facts about reasons for action. At this point, the reader is invited to score Kant’s explanation accordingly.

Perhaps surprisingly, Scanlon’s theory of reasons for action is also well understood as a version of constitutivism that explains facts about reasons for action in terms of the goodness-fixing kind deliberator (Scanlon 1998). What is Scanlon’s preferred account of what it is to deliberate well? In Scanlon’s view, there is a class of judgment-sensitive attitudes, where the distinctive feature of this class is that the attitudes are acquired and lost depending on whether the person who has them takes there to be sufficient reasons for acquiring and losing them. Beliefs are judgment-sensitive attitudes par excellence, and intentions and intrinsic desires in the very broad sense in which we have been understanding them here are also such attitudes. To deliberate well, according to Scanlon, is therefore just to acquire and lose judgment-sensitive attitudes—that is, beliefs, intentions, and intrinsic desires—as a result of a sensitivity to the reasons for them.

The reason relation, as Scanlon understands it, is a primitive four-place relation that relates a fact, a person, a circumstance, and a judgment-sensitive attitude: the fact provides the person in that circumstances with a reason to have that judgment-sensitive attitude (Scanlon 2014). He may suppose that, in a situation of conflict, there is a certain fact—perhaps the fact that certain actions would accord with principles for the general regulation of society that no one could reasonably reject—that provides the individuals in those circumstances with a reason to intend to perform actions of that kind. If such principles captured the fair terms of cooperation, then the individuals in those circumstances would each have reasons to intend to cooperate on fair terms. They would have these reasons no matter what species of living thing they are; no matter what intrinsic desires they would happen to have after they exercised their imaginations, given their embodiment, their upbringing, and their culture; and no matter what intrinsic desires they would end up with after having thoughts about what’s required for motivations to conform to universal laws.

Note that what we have here is an explanation of which judgment-sensitive attitudes agents in such circumstances have reason to have, not what they have reasons to do. This is why we do well to understand Scanlon’s explanation of reasons for action in constitutivist terms, as it is constitutivism that allows us to turn reasons for judgment-sensitive attitudes into reasons for action. What agents have reason to do is explained by what they would be motivated to if they deliberated well, so if people deliberate well to the extent that they are sensitive to the independently given reasons for judgment-sensitive attitudes, and badly to the extent that lack such a sensitivity, then if there are reasons for those in circumstances of conflict to intend to cooperate on fair terms, it follows that they have reasons to cooperate on fair terms. They have such reasons because this is what they would be motivated to do if they were to deliberate well.

As I said, Scanlon seems to commit himself to constitutivism of this kind when he tells us that “reason for action’ is not to be contrasted with ‘reason for intending’” (Scanlon 1998, p. 21). His idea seems to be that since what makes an event an action is the fact that it was the product of a certain judgment-sensitive attitude, namely an intention to perform the action, so what makes an action one that an agent has reason to perform
is the fact there are reasons for the agent to have the intention to perform it. We get constitutivism by restating this as a counterfactual. What makes an action one that the agent has reason to perform is the fact that the agent would intend to perform it if he were sensitive to the reasons for having intentions.

Is Scanlon’s explanation of reasons for action, so understood, a reductive or a non-reductive version of constitutivism? Since it appeals to a normative feature in explaining the constitutive features of deliberators—the primitive four-place reason relation—that cannot be explained in non-normative terms, it is a non-reductive version. Deliberators are those who have the capacity to be sensitive to reasons for judgment-sensitive attitudes in the formation of such attitudes. But since this normative feature is a distinct normative feature from the normative feature that his constitutivism explains—namely, reasons for action—it follows that the requisite explanatory relations between facts about reasons for action and facts about the constitutive features of deliberators remain intact. Facts about reasons for judgment-sensitive attitudes explain facts about the constitutive features of deliberators, and facts about the constitutive features of deliberators explain facts about reasons for action. At this point, the reader is invited to score Scanlon’s explanation.

Though we have considered various versions of constitutivism about reasons for action, we have not yet discussed two of the better-known versions of the view. Christine Korsgaard and David Velleman both defend versions of constitutivism, but they conceive of the view very differently from the way it has been characterized here (Korsgaard 1996, 2009; Velleman 1989, 1996). As they see things, constitutivism is the view that reasons for action are to be explained in terms of the constitutive aim of agents. In Korsgaard’s view, this is the aim of self-constitution. In Velleman’s view, it is the aim of making sense. According to some who follow Korsgaard’s and Velleman’s lead, the constitutive aims are something else again: Paul Katsanafas, for example, thinks that the constitutive aim of agents is some end or other plus the aim of encountering and overcoming obstacles or resistances in the pursuit of that end (Katsafanas 2013, forthcoming).

My reason for not discussing these views until now is because their way of understanding constitutivism is so unhelpful. The claim that agents have a constitutive aim is either false or metaphorical. Suppose we understand the claim non-metaphorically, and a constitutive aim is some sort of intrinsic desire or intention that every agent is supposed to have simply by virtue of being an agent. The problem with this is that there are evidently agents—many non-human animals, human infants, children, and those who are cognitively impaired—who lack the conceptual sophistication required to have any of these aims. More generally, though it seems right that agents may have some intrinsic desires or intentions or other to count as agents, which desires or intentions they have seems to be an entirely contingent matter. In other words, in the non-metaphorical sense, agents do not have constitutive aims.

This leaves open the possibility that talk of agents having a constitutive aim is metaphorical, but then the question is how the metaphor is to be understood in non-metaphorical terms. The best way of understanding it, in my view, is to suppose that the kind agent is a goodness-fixing kind: for Korsgaard, to be an agent is to possess the capacity to constitute oneself as a unified agent; for Velleman, it is to possess the capacity to make oneself make sense; and for Katsanafas, it is to have the capacity to have ends and to encounter and overcome obstacles in the pursuit of those ends. Though each of
these ways of understanding their views is more plausible than the claim that agents have the corresponding constitutive aims, the elaboration and evaluation of their views so understood would take us some distance from anything that these theorists actually say in defense of their views. It therefore seems best to leave a discussion of them for another occasion.

The main objection to constitutivism is the well-known agency-shmagency objection due to David Enoch (2006). Since the objection was originally stated as an objection to the constitutive aim theorists’ versions of constitutivism, it will be helpful to restate it as an objection to the more plausible versions of constitutivism as the view has been understood here. These are the versions of constitutivism that explain reasons for action in terms of the constitutive features of deliberators, where deliberation is understood in Williams’s or Kant’s or Scanlon’s way. The agency-shmagency objection is that it is opaque why reasons, understood in any of these ways, have the authority that reasons are supposed to have.

Imagine people who are motivated to act in the way that some guru instructs people to act, and imagine that what the guru instructs people to do is completely different from what they would be motivated to do if they deliberated well. According to all of these versions of constitutivism, these people act in ways they have no reasons to act. They are deliberators, so what they have reason to do is fixed by what they would be motivated to do if they deliberated well. But, the objection goes, what’s so special about the goodness-fixing kind deliberator? We could coin the word ‘shmeliberator’ to pick out those who could act in the way the guru instructs people to act, and, on plausible assumptions, all and only deliberators are shmeliberators and vice-versa. Shmeliberator, like deliberator, is therefore a goodness-fixing kind—we can order people from best to worst according to how well their acts accord with the guru’s instructions—and we can define ‘shmeasons for action’ as what people would be motivated to do if they were excellent shmeliberators. If we deliberate well, we shmeliberate badly, and if we shmeliberate well, we deliberate badly. So, what makes it the case that the considerations that would motivate us if we deliberated well have authority over us in a way that those that would motivate us if we shmeliberated well don’t? The agency-shmagency objection is that constitutivism can give no good answer to this question.

There have been many replies to the agency-shmagency objection, and many replies to these replies (examples include Ferrero 2009; Tiffany 2012; and Silverstein 2015). The best reply, in my view, focuses on what it is for considerations that motivate us to be authoritative. On one view, a consideration that motivates us is authoritative if it is a consideration that motivates an excellent member of some goodness-fixing kind of which we are members. On this way of understanding authority, the considerations that motivate shmeliberators are indeed authoritative, so there is nothing for constitutivists to explain. On an alternative view, the considerations that motivate us are authoritative only if they are considerations that motivate us to do what we have reason to do. On this way of understanding authority, constitutivists do need to explain why the considerations that motivate shmeliberators are not authoritative, but this isn’t a problem, as they have a good explanation of this.

Why suppose that we do not have reason to what we would be motivated to do if we shmeliberated well? The answer, according to the constitutivists we have considered, is the intrinsic plausibility of supposing that we find out what we have reason to do by delib-
erating, where deliberation is understood in Williams’s or Kant’s or Scanlon’s way, and the
fact that what we ordinarily take to be the extension of what we have reason to do is well
captured by what we would be motivated to do if we deliberated well, where deliberation
is understood in one of these ways. By contrast, it is completely implausible to suppose
that we find out what we have reason to do by finding out whether what we are motivated
to do corresponds to the guru’s instructions, and the extension of what we would be
motivated to do if we deliberated well is therefore a poor fit with what we ordinarily
take to be the extension of what we have reason to do.

Of course, those who put forward the agency-shmagency objection might think that
no one can explain what it is to have a reason for action, as such facts are explanatorily
basic (see Wiland’s comparison of the agency-shmagency objection to the open-question
argument [2012], and McPherson’s discussion of robust versus formal normativity
[2011]). Though they might agree that we find out what we have reason to do by delib-
erating, they might think that deliberation is to be understood as the exercise of the
capacity to detect instantiations of the primitive property of an act’s being something
there is reason to do. This primitivist view of reasons for action is indeed a competitor to
constitutivism. However, to the extent that we think we can give an adequate description
of deliberation without assuming this primitivist view of reasons for action, perhaps a
description like Williams’s or Kant’s or Scanlon’s, or perhaps a description in some other
terms entirely, we should reject it in favor of constitutivism.

ACKNOWLEDGMENTS

Thanks to Jamie Fritz, Tristram McPherson, and David Plunkett for their very helpful comments on a draft.

REFERENCES


Habermas, Jurgen 1996: Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy,


University Press).

and Rationality (Oxford: Oxford University Press).


Volume (63) pp. 113–137.


Scanlon, Thomas M. 1998: What We Owe To Each Other (Cambridge, MA: Harvard University Press).