



Reid on the Autonomy of Ethics: From Active Power to Moral Nonnaturalism

ABSTRACT: *Thomas Reid has the unusual distinction of arriving at a metaethical position very much like G. E. Moore's via a route very similar to that employed by the Kantians. That is, Reid embraces a version of nonnaturalist moral realism by appeal not to open question-style considerations but to a particular account of agency. In this essay, we reconstruct Reid's agency-centered argument for his constitutivist version of moral nonnaturalism, highlighting its commitments. Having presented Reid's argument, we close by considering a prominent contemporary Kantian view, namely, Christine Korsgaard's, and identifying where, despite their common commitments, Reid and Korsgaard part company. The comparison, we suggest, is instructive because it allows us to see more clearly why the link between agency-centered approaches to ethical theorizing and nonrealist, constitutivist views of morality, such as Korsgaard's, is deeply contingent.*

KEYWORDS: history of ethics, agency, action, metaethics, Reid, active power, moral nonnaturalism

The primary question that concerns G. E. Moore at the outset of the *Principia Ethica* ([1903] 1993) is whether or not moral properties are natural—whether they are the sorts of thing investigated by, or that play an explanatory role in, the natural sciences. Those familiar with Moore's answer to his question know that he arrives at it by employing an approach that consists in reflecting on the meaning of ethical terms and drawing conclusions about the nature of moral reality on the basis of his findings. In so doing, Moore employs what we might call a *content-centered* approach to settling fundamental metaethical questions. (We are using the term 'content' broadly to stand for what some would call the object of belief.) Moore's content-centered approach makes very few presuppositions about moral agency beyond the fact that we can grasp the meaning of moral terms to some sufficient degree.

Some have concluded that Moore's approach to settling fundamental metaethical questions is a dead end. A good many of these philosophers take their cues from Kant, suggesting that when trying to settle fundamental metaethical questions,

Our thanks to Tyler Doggett, Esther Kroeker, Andrew Malcovsky, Douglas McDermid, René van Woudenberg, the editor of this journal, and several anonymous referees for comments or discussion regarding earlier versions of this essay.

we should reflect not on the meaning of moral terms but on the nature of moral agency, determining what is required to be such an agent. These philosophers employ what we will call an *agency-centered* approach. Nearly without exception, these philosophers accept some version of constitutivism, claiming that there are certain normative principles upon which someone must act in order to be an agent or for her behavior to count as an action (see Korsgaard 2009, Velleman 2009, Walden 2012, and Katsafanas 2013). And nearly without exception, they present constitutivism as a nonrealist rival to realist views such as Moore's.

Thomas Reid has the unusual distinction of arriving at a position very much like Moore's via a route very similar to that employed by the Kantians. To use the terminology just introduced, rather than employ a content-centered approach, Reid employs an agency-centered approach to defend a version of nonnaturalist moral realism that draws upon his agent-causal theory of action.¹ Those familiar with work of the so-called British moralists know that Reid was not alone in combining an agent-causal view with moral nonnaturalism; Samuel Clarke and Richard Price did as well (see Clarke [1998] and Price [1787/1948]). But Reid is distinctive in arguing for his nonnaturalist realist view by appealing to his favored account of agency. In what follows, we reconstruct Reid's agency-centered argument for nonnaturalist moral realism, highlighting its commitments.

Our aim in doing so is twofold. On the one hand, we are interested in contributing to and expanding the contemporary discussion of the British moralists (see Darwall 1995, Gill 2006, and Harris 2005). Since Reid's metaethical views have received only cursory attention in this discussion, we dedicate a significant portion of our discussion to presenting how Reid's views on agency drive his metaethical commitments. On the other hand, we want to present a line of argument and a position that are themselves of philosophical interest, despite having been largely overlooked in the literature. As will be evident, our aim in presenting this line of argument and position is not so much to defend them but to highlight the distinctive contribution they make—and perhaps could continue to make with further supplementation and defense—to our understanding of metaethics. In the last section of the paper, we compare Reid's views to those of Christine Korsgaard, identifying where Reid and Korsgaard part company despite their considerable common commitments. The comparison proves to be instructive because Korsgaard and others present agency-centered constitutivism as naturally aligned with naturalistic metaethical nonrealism and, thus, as an alternative to positions such as Reid's (see Walden [2012], for example). Having Reid's views before us helps us to see that there is no such natural alignment, as agency-centered constitutivism is compatible with nonnaturalist moral realism.

¹ Reid does, however, have his Moorean moments; see Reid [1788] 2010: III.ii.v: 129; abbreviated hereinafter as EAP. In what follows, we will often refer to Reid's nonnaturalist moral realism as simply 'nonnaturalism' or 'realism'. So, when we contrast Reid's position with nonrealist positions such as Korsgaard's, the contrast is between nonnaturalist moral realism and moral nonrealism. In calling Reid's view a version of nonnaturalism, we are not using Reid's own terminology, but terminology made popular by Moore to refer to a type of view that Reid, Moore, and others accept.

I. Ethics is Autonomous

The argument that Reid offers for his broadly nonnaturalist position stretches over most of the *Essays on the Active Powers of Man*, taking various detours along the way. We propose to work backward to forward by first presenting the conclusion that Reid wishes to establish and then offering a formalization of his argument for it. Having offered this formalization, we will then consider his reasons for accepting its key premises.

The conclusion Reid wishes to defend is a thesis we will call:

Ethics is Autonomous: the findings and methods of the natural sciences have limited implications for moral knowledge and deliberation. Employing the methods of the natural sciences does not yield moral knowledge. Any finding of these sciences has ethical significance only insofar as it can become the content of ethical deliberation.

In order to argue for Ethics is Autonomous, Reid first argues for a related thesis that we will call:

Moral Nonnaturalism: moral features are fitted neither to be the subject matter of the natural sciences nor to play an explanatory role in these sciences. In a phrase, these features are not ‘taxonomic’ in the natural sciences.²

In [section 4](#), we will adduce textual evidence for attributing these two theses to Reid. In the meanwhile, let us make two initial observations about them.

First, while these two claims are not identical, they are intimately connected. Under the assumption that moral knowledge concerns moral features, then Moral Nonnaturalism implies Ethics is Autonomous (an implication we will exploit later). Second, when Reid tells us that moral features are not—to use the terminology introduced above—taxonomic in the natural sciences (and, hence, not ‘natural’), he has something fairly specific in mind. In Reid’s view, to say that moral features are not taxonomic in the sciences is to claim that these features do not constitute, nor are they referred to by, the fundamental natural laws that govern the happenings in the world as revealed to us by Newtonian physics. It is worth emphasizing that Reid’s endorsement of Moral Nonnaturalism does not stem from a general discontent with or skepticism about the natural sciences. To the contrary, Reid was a devout follower of Newton and regularly emphasized the power of Newton’s findings and methodology (cf. EAP: IV.ix: 251; Callergård 2010 and Wood 2004). But Reid was also keen to insist on the limitations of Newton’s findings and methodology. One of these limitations, Reid claims, is that Newton’s findings and methodology do not apply to the normative domains. (We will touch on this claim again later.)

² This is a fairly standard characterization of nonnaturalism; see Moore (1903/1993), Wiggins (1993), and Enoch (2011:103) although Enoch rightly notes that it requires further amplification and nuance. In what follows, we use the term ‘moral feature’ broadly so that it pertains to moral properties, facts, and truths. We borrow the term ‘taxonomic’ from Fodor (1991).

These two points having been made, let us now present an argument that captures the pattern of thought that leads Reid to endorse Ethics is Autonomous. The argument runs thus:

- (1) Human persons are agents.
- (2) To be an agent is to be the sort of being who acts, where acting consists in exerting active power to actualize some behavior-state in light of an agent-reason.
- (3) So, human persons (often) act in light of agent-reasons.
- (4) Some of the agent-reasons in light of which human persons act (including those by which they regulate their behavior) are moral.
- (5) The moral reasons in light of which human persons act (including those by which they regulate their behavior) are either natural or not.
- (6) Moral Nonnaturalism: these reasons are not natural. For if they were, then they would be taxonomic in the sciences, which they are not.
- (7) Since Moral Nonnaturalism, therefore Ethics is Autonomous.

In what follows, we will not discuss premises (1) or (5), since these should be acceptable to most parties to this discussion. Instead, we will focus on the argument's remaining three premises: (2), (4), and (6). Premise (2) states Reid's account of agency. Since Reid's views on this matter are complex, this section will occupy much of our attention. With his account of agency in hand, we will then turn to Reid's defense of premise (4). What will emerge from our discussion is that Reid does not just argue that (4) is true; he also holds that (4) is an implication of a proper understanding of practical rationality according to which practical rationality is exercised in the regulation of behavior. Finally, we will move to (6), which we can treat more briefly given what we have said about Reid's understanding of agency and practical rationality.

2. Premise (2): The Dual-factor Theory of Agency

Imagine yourself standing on a city street corner watching someone vigorously wave his hand above his head. Determining that he is engaged in behavior of some sort is easy: you only have to look at his bodily movements. In contrast, determining whether he is performing an action—whether it be flagging a taxi cab or dancing—is more difficult. What, after all, would explain the fact that his behavior counts as an action? (We restrict our attention here to free human actions; as we explain below, Reid allows that there is a loose and nonphilosophical sense in which humans and nonhuman animals are capable of acting, albeit not freely.)

This question—sometimes referred to as *the action problem*—is one to which philosophers offer widely different and rival answers. Reid's answer to the action problem can be stated crisply: an item of behavior counts as an agent's action just in case and because that agent is its efficient cause (EAP I.i: 13). Lying behind this

formulation of Reid's view is a complex position that Reid defends at length. To understand Reid's reasons for accepting premise (2) of his argument, we will need to have its main lines before us.

Begin by distinguishing an *efficient cause* from an *occasional cause* (we are drawing on Reid's terminology in EAP I.vii). To say that something is an efficient cause, according to Reid, is to claim that it is an entity endowed with active power—what many today would call an *agent cause*. In contrast, to say that something is an occasional cause—or, as we will also say, a *nomic cause*—is to claim that it is an event-type that is linked in a law-like manner to other event-types of some range. Reid says rather less about what exactly renders something an efficient cause (that is, an agent cause) than one would hope, largely because he claims that he can offer no definition of power and, hence, of active power. ('Power is a thing so much of its own kind, and so simple in its nature, as not to admit of a logical definition' [EAP: I.i: 7].) Still, he makes several observations about active power that are important for our purposes.

First, active power is what is exerted in action: 'The exertion of active power we call *action*' (EAP I.i: 13). Second, active power necessarily comes with options: 'Power to produce an effect implies power not to produce it' (EAP I.v: 29). That is, if an agent can actualize a behavior-state through the exertion of active power, then the agent can also not actualize that same behavior-state through the exertion of active power. And, third, for any action performed by an agent (to which that agent attaches any importance), there must be something in the agent's state of mind preceding that action that inclines him toward or recommends that determination (EAP II.i: 51). That which inclines an agent to act or recommends that he act is what Reid calls a 'principle of action' or a 'motive'³

Those familiar with Reid's discussion of motives know that it presents the reader with several interpretational challenges, since Reid uses the term 'motive' in a variety of incompatible ways. For our purposes, it will not be necessary to settle these exegetical issues (we discuss this point in greater detail in Cuneo and Harp 2016). It is enough to see that Reid operates with a particular theory of action, which we will call his *dual-factor* theory of action.

The key to understanding Reid's dual-factor theory is to recognize that Reid uses the term 'end' in two different ways. At some points, Reid uses the term to refer to what favors the performance of an action—that is, that feature of a situation that makes that action worth performing or recommends its performance. Call ends of this sort *agent-reasons*. Reid tells us that there are only two types of fundamental agent-reasons: an agent's 'good on the whole' (or well-being), and 'duty' (EAP III.iii.i: 154). These agent-reasons are what Reid calls the 'rational principles of action', and they are the first factor in Reid's dual-factor theory. At other points, Reid uses the term 'end' to refer to those states that an agent can endeavor to bring

³ For Reid, 'principle of action' is the general category: for every action that an agent performs, there must be some corresponding principle of action that leads that agent to perform that action. In this general category, Reid contrasts motives (both animal and rational motives) with mechanical principles of action such as instincts and habits, and uses the more specific term 'rational motive' to stand for what we later call 'agent-reasons'. Because we are not concerned with mechanical principles of action here, we will use the terms 'principle of action' and 'motive' interchangeably, keeping in mind that motives need not be rational motives.

about or actualize through the exertion of his active power. Reid sometimes calls these ends ‘actions’, but to avoid confusion, we will call them *end-states* (see EAP II.i). Examples of end-states would be your kicking a ball into the left-hand corner of a soccer goal or pruning an apple tree. End-states are the second factor in Reid’s dual-factor theory.

Fundamental to Reid’s account of action is the claim that we need to understand the exertion of active power and, hence, action in light of ends of both sorts. For in the paradigmatic case, when an agent acts, that consideration an agent views as rendering his action worth pursuing—what functions as ‘advice’ for an agent—is not identical with the end-state he endeavors to bring about (EAP II.ii: 59). When you prune an apple tree, for example, there is typically some consideration distinct from this state that recommends doing so—perhaps in doing so you follow an arborist’s advice. If this is right, we can formulate Reid’s dual-factor theory as follows:

Cases of indifferent action aside, an agent *S* acts for the sake of an end if and only if there is some end-state *E* such that *S* actualizes *E* through the exertion of his active power in light of some agent-reason *R*.

The dual-factor theory of action captures what is arguably the core of Reid’s account of agency and action. But we do not yet have any feel for why Reid accepts the dual-factor theory. As it happens, Reid offers a battery of arguments in favor of it—arguments that charge the theory’s main rivals with offering a vacuous account of motivation, not being able to explain action taken out of indifference, and not adequately explaining how ascriptions of responsibility could make sense if their view were true (see EAP IV). Rather than attempt to summarize these arguments, we propose to present only one of them. Our aim in doing so is not to dig into its finer details, but to extract two constraints on a satisfactory theory of action that Reid accepts. Doing so will not only help us to see what motivates Reid’s agency-centered approach to ethical theorizing, but it will also throw the comparison between Reid’s view and Korsgaard’s Kantian position into sharper relief.

First a bit of stage-setting: Reid fashions the dual-factor theory against a position that he calls the system of necessity (EAP IV.i: 198). Advocates of the system of necessity do not deny that agents act, and indeed often insist that agents act freely. Still, these philosophers understand action much differently from Reid. Perhaps the best way to see how they differ is to call to mind Reid’s distinction between efficient causes (that is, agent causes) and occasional causes (that is, nomic causes). Advocates of the system of necessity—among whose number Reid includes Spinoza, Hobbes, Hume, Hartley, and Priestley—think of action entirely in terms of occasional causes. Actions, these philosophers claim, are entirely determined by an agent’s motives, which (they maintain) are not ends but mental states and events.⁴

⁴ That Reid held this to be the crucial issue is evident in the following passage: ‘Since Mr Collins wrote his Essay on Liberty I think the Controversy has turned chiefly upon the Influence of Motives as we may see in the writings of Dr Clark of David Hume, of Lord Kames, of Priestly & those that opposed them upon this point’ (cited by the editors of EAP on p. 213, n.17).

According to the system of necessity, a piece of behavior counts as an agent's action just in case and because it is caused in the right way by one or another of that agent's mental states. Reid, as we have seen, rejects this approach. According to Reid, in the paradigm case, some item of behavior counts as an agent's action just in case and because the agent is that action's efficient cause (that is, agent cause), and the exertion of her active power is for the sake of actualizing some end-state in light of some agent-reason.

With that noted, let us turn to Reid's argument, wherein Reid asks us to consider a complex action sequence, such as one in which an agent who exhibits practical rationality deliberates and takes certain means to bring about some end she aims at. Imagine, for example, that an agent is deliberating about whether to go on a hike. This agent might have incentives of various sorts, such as desires (say, to get exercise), aversions (say, to heights), and intentions to actualize certain end-states (such as enjoying autumn foliage). If this agent exhibits practical rationality, then when deliberating and acting, she must intervene in ways that are appropriate: for example, by considering certain options, but not others; resisting certain desires, but endorsing others; forming and enacting certain intentions, but not others, and so on. That is, if this agent exhibits practical rationality, she must form auxiliary intentions, such as paying attention to the autumn foliage, and must resist doing such things as making plans that conflict with taking a hike. None of these activities, however, is performed by mental states themselves. Intentions, for example, are not self-enacting; rather, agents must enact them. And deliberations are not resolved merely through the desires of the agent; agents must resolve the deliberations. If this is so, Reid claims, advocates of the system of necessity fail to recognize and account for important aspects of the agent's role in acting, and fail to include them in their account of agency.

Here is how Reid puts the point in his own words. Referring to the person of 'wisdom and understanding', Reid writes:

But if all his particular determinations, which concurred in the execution of this plan, were produced, not by himself, but by some cause acting necessarily upon him, there is no evidence left that he contrived this plan, or that he ever spent a thought about it.... If it be said, that all this course of determinations was produced by motives; motives surely have not understanding to conceive a plan, and intend its execution. We must therefore go back beyond motives to some intelligent being who had the power of arranging those motives, and applying them, in their proper order and season, so as to bring about the end. (EAP IV.viii: 241)

This argument will be familiar to those immersed in contemporary debates regarding the nature of agency; Reid is criticizing the system of necessity with a variant of the 'disappearing agent' objection (cf. Velleman 1992; see also the discussion in Yaffe 2004).

Reid's version of the objection turns on the claim that any satisfactory theory of action must satisfy the following two constraints. The first articulates the following general criterion:

When an agent's actions are practically rational, she must make interventions of the appropriate kind in the causal process leading to them.

The second specifies what it is for an intervention to be of the appropriate kind:

For an intervention to be of the appropriate kind, then it must be produced, not merely *by* the agent (for example through motives), but through an *exercise of agency*.

Reid's charge is that the system of necessity fails both constraints. In the passage quoted above, Reid argues that motives themselves cannot make interventions of the appropriate kind in action. However, according to the system of necessity, there is nothing an agent can be except a collection of motives: it is motives that cause action, fully determining what an agent does (EAP IV.iv: 213). It follows that, according to the system of necessity, agents cannot make interventions of the appropriate kind in the causal process leading to actions, and so the position fails the first constraint. As for the second constraint, since agents cannot make interventions of the appropriate kind, a fortiori they cannot intervene through an exercise of agency. In contrast, the dual-factor theory, which maintains that agents can do such things as arrange motives and conceive of and implement plans by the exercise of active power, satisfies both constraints.

We can now see how Reid's case for (2) is supposed to run. Reid is assuming that there are two constraints on a satisfactory theory of action explanation. The dual-factor theory, Reid maintains, straightforwardly satisfies these constraints. Its most viable rival, the system of necessity, cannot satisfy either constraint. Reid takes this to be an excellent reason to endorse (2).

Our aim on this occasion is not to determine whether Reid is right about this, let alone explore ways in which his argument might be buttressed or sharpened. We will, however, say two things on its behalf. First, while we have spoken of Reid's dual-factor theory and the system of necessity as individual theories, they are probably better thought of as families of theories—indeed, as families into which we can fit a good many (and perhaps most) contemporary theories of action. If this is right, then Reid's argument is directed toward not two individual theories of action that happened to be prevalent in his day but toward a wide range of such theories that have currency. Second, while the philosophical landscape has changed significantly since Reid offered his defense of the dual-factor theory, the two constraints that Reid identifies are widely accepted among contemporary theories of action. In this respect, Reid had his eye on considerations that such theories have considered important to address. If the dual-factor theory satisfies

these two constraints, then Reid was right to conclude that that is a consideration in its favor, even if it is not decisive.

3. Premise (4): From Practical Rationality to Moral Reasons

When Reid presents his account of agency, it is evident that he believes that discussions of agency often gravitate toward one of two extremes. At one extreme is the system of necessity, which, it appears, eliminates the role of the agent in action explanation, attributing to mental states activities that could only be rightly attributed to agents. This view fails to illuminate how actions are, in any interesting sense, ‘up to’ the agent. At the other end of the spectrum are views that emphasize acting out of pure indifference. According to these positions, the indifferent agent is the free agent par excellence, the one who rises above any inclinations she may have, simply choosing to act one way or another. In doing so, this agent exhibits what has often been called ‘liberty of indifference’. While Reid believes that there can be ‘motiveless’ actions and draws upon this point to attack the system of necessity, he does not accord liberty of indifference the sort of importance that other figures do, such as William King and Edmund Law (see Harris [2005: ch. 2]; for Reid’s defense of the possibility of motiveless actions, see EAP: IV.iv). Instead, Reid attempts to strike a mean between these two extremes, claiming that in the paradigm case we are neither simply pushed around by what he calls ‘incitements’ nor indifferent to them. How does Reid propose to do that?

Reid’s answer to this question is found in a series of connections he identifies between active power, rationality, and the principles of action. A being endowed with active power, Reid says, must have practical reason (EAP III.iii.i). But to have practical reason, Reid continues, is to be capable of being governed by law:

A being endowed with the animal principles of action only, may be capable of being trained to certain purposes by discipline, as we see many brute animals are, but would be altogether incapable of being governed by law.

The subject of law must have the conception of a general rule of conduct, which, without some degree of reason, he cannot have. He must likewise have a sufficient inducement to obey the law, even when his strongest animal desires draw him the contrary way.

This inducement may be a sense of interest, or a sense of duty, or both concurring. These are the only principles I am able to conceive, which can reasonably induce a man to regulate all his actions according to a certain general rule or law. They may therefore be justly called the *rational* principles of action, since they can have no place but in a being endowed with reason, and since it is by them only, that man is capable either of political or of moral government.

Without them human life would be like a ship at sea without hands, left to be carried by winds and tides as they happen. (EAP III.iii.v: 168, emphasis in original)

This passage is instructive for a pair of reasons. We begin with the first.

Recall that Reid says that the rational principles of action come in two varieties: those that concern one's good on the whole and those that concern duty. In principle, one could imagine beings who act entirely in terms of the former principle: they act for prudential but not moral considerations. Given a proper understanding of one's good on the whole, however, Reid indicates that this is impossible:

The right application of [the principle of our good upon the whole] to our conduct requires an extensive prospect of human life, and a correct judgment and estimate of its goods and evils, with respect to their intrinsic worth and dignity, their constancy and duration, and their attainableness. He must be a wise man indeed, if any such man there be, who can perceive, in every instance, or even in every important instance what is best for him upon the whole, if he have no other rule to direct his conduct.

However, according to the best judgment which wise men have been able to form, this principle leads to the practice of every virtue. It leads directly to the virtues of prudence, temperance and fortitude. And, when we consider ourselves as social creatures, whose happiness or misery is very much connected with that of our fellow-men; when we consider, that there are many benevolent affections planted in our constitution, whose exertions make a capital part of our good and enjoyment; from these considerations, the principle leads us also, though more indirectly, to the practice of justice, humanity, and all the social virtues. (EAP III.iii.iii: 163–64; cf. III.iii.iii: 159)

Reid uses fairly weak language here to describe the relation between having an accurate conception of one's good on the whole and moral virtue, describing the first as 'directly leading' to the latter. But consider what Reid actually says: determining one's good on the whole requires 'a correct judgment and estimate of its goods and evils, with respect to their intrinsic worth and dignity'. It also must involve a considerable measure of self-control, negotiating 'temptations that lie in his way' (EAP III.iii.iv: 165). Indeed, Reid elsewhere supplements what he says here with stronger language. Regarding our good on the whole, Reid writes: 'this principle is so similar to the moral principle... and so interwoven with it, that both are commonly comprehended under the name of *reason*' (EAP III.iii.ii: 159, emphasis in original).

A closer look at what Reid says, then, reveals that he has in mind a much tighter connection between one's good on the whole and moral virtue than some of his language would indicate. Specifically, Reid holds that having a reasonably accurate take on one's good on the whole requires exercising traits such as prudence, temperance, and fortitude, which Reid takes to be moral virtues. In characterizing these traits as moral, Reid appears to be following Butler's thinking, holding both that proper care of self is an ethical matter and that being properly related to others is constitutive of an agent's good on the whole (McNaughton [2013: 385] highlights

this dimension of Butler's thought. In Butler's view, the fact that we hold each other accountable and even blame each other for failing to care for self is excellent evidence that care of self falls within the ethical domain.) The passage quoted above, then, is one in which Reid points to a case of 'normative entanglement', indicating that a proper understanding of the concepts that belong to one normative domain (such as prudence) requires adverting to concepts that belong to another such domain (in this case, morality). Given the assumption that there are features that satisfy these concepts, it follows that there are moral features, a claim that lies at the core of Reid's realism (see Cuneo 2004).

We now turn to the second reason why the first passage we have quoted is instructive, which is that it highlights Reid's conviction that the principle of duty plays distinctive roles in action, these roles being both normative and psychological. (We follow Reid's terminology in calling the second rational principle of action the principle of duty, though the terminology is misleading; the phrase 'principle of duty' is really just a placeholder for a host of more specific moral principles, many of which do not directly concern moral obligation but virtues and rights.)

The moral principles of action play a distinctively normative role, according to Reid, because they are authoritative: 'all wisdom and virtue consist in following [their] dictates' (EAP II.iii: 58). Their authoritative character is particularly manifest in deliberative contexts when some desire presents itself to the mind. When it does so, Reid stresses, you have the ability to distance yourself from it, asking: Ought I to act on this desire or impulse? Would it contribute to my good on the whole? And does duty permit acting upon it? If you determine that the rational principles of action prohibit acting on the desire, the matter is settled; these principles have, so to speak, final say on the matter—although, of course, you might not heed their authority.

In addition to playing this distinctively normative role, the rational principles of action play a psychological role since they can be internalized by an agent and an agent can regulate her behavior by appeal to them. When expanding on this point, Reid sometimes speaks in a Kantian vein, adverting to our ability not simply to appeal to law but also to 'give law' to ourselves (EAP II.ii: 57). When he does speak this way, Reid has in mind our ability to form a 'fixed purpose, or determination, to act according to the rules of justice' (EAP II.iii: 66) or 'to do good when we have opportunity, from a conviction that it is right, and is our duty' (EAP II.iii: 67). This fixed resolution, says Reid, is nothing other than virtue (EAP II.iii: 66).

The upshot is twofold: first, the rational principles of action are not merely agent-reasons, i.e., those considerations in light of which an agent acts or which render an act worth performing or otherwise attractive in the eyes of an agent. They are also authoritative considerations by appeal to which an agent can *regulate* her behavior, engaging in self-governance. Second, this self-regulation often consists in giving oneself a law, laying down a resolution to behave in certain ways and not others. When we do so, and reliably act in accordance with this resolution, Reid says that we exhibit virtue—virtue, for Reid, having its own distinctive regulative function, this regulative function being itself the product of an agent's exertion of her own active power.

If this is so, we can now see why Reid accepts:

- (4) Some of the agent-reasons in light of which human persons act (including those by which they regulate their behavior) are moral.

The argument for (4) runs thus: we are, according to premises (1) and (2), endowed with active power. And if we are, then we can act for ends or reasons, which implies that we have practical reason. But to have practical reason is to be such that one can act from the rational principles of action, being 'governed by law'. In principle, we could be motivated to act purely out of prudential concerns. But on the supposition that we sometimes understand our good on the whole more or less correctly, we must also act in light of moral considerations. In a large range of cases, however, to act in light of moral considerations is not simply to appeal to these considerations; it is also to be governed by them so that they regulate one's behavior. That, among other things, consists in internalizing these considerations in such a way that one acts out of virtue, which is simply the fixed resolution and the tendency to act for moral reasons of a certain range.

4. Premise (6): Nonnaturalism

Reid's dual-factor theory of action tells us that, in the paradigm case, an agent acts through the exertion of her active power to bring about some behavior-state in light of an agent-reason. We have explored (at least) one reason why Reid thinks that some of these reasons are moral: they are among the considerations by which agents regulate their behavior when pursuing their overall good such that failure to do so renders these agents liable to reactions such as blame and admonition. We can, however, ask about the character of both behavior-states and agent-reasons. Are they, to use the terminology employed earlier, natural since they are taxonomic in the natural sciences?

The following passage contains Reid's most explicit answer to these questions:

There are many important branches of human knowledge, to which Sir Isaac Newton's rules of Philosophizing have no relation, and to which they can with no propriety be applied. Such are Morals, Jurisprudence, Natural Theology, and the abstract Sciences of Mathematicks and Metaphysics; because in none of those Sciences do we investigate the physical laws of Nature. There is therefore no reason to regret that these branches of knowledge have been pursued without regard to them. (Reid 1995: 186)

Reid offers a variety of considerations for the verdict he offers here, beginning with the observation that while moral principles and natural laws are both 'properly called *laws of nature*', they are of different kinds. The 'moral laws of nature... are the rules which God has prescribed to his rational creatures for their conduct' while 'the physical laws of nature are the rules according to which the Deity commonly acts in his natural government of the world' (EAP IV.ix: 251). More specifically, laws

of each type have different modal profiles: while the laws of nature are contingent, the fundamental moral principles (or laws) are necessary (cf. Reid [1785] 2002: Vi.vi: 494–95; hereinafter abbreviated as EIP). Furthermore, laws of each type have different epistemic profiles: while natural science proceeds ‘from particular facts in the material World, to collect by just Induction the Laws that are less general’ (Reid 2002: 142), the fundamental moral laws are self-evident, grasped by rational intuition (EAP V.i: 272–76).

It is in his chapter ‘Of the Influence of Motives’, however, that Reid presents his most vivid and developed case for the claim that moral features are not taxonomic. At the heart of Reid’s discussion is the claim that, properly understood, motives are ends, being either agent-reasons (that in light of which agents act) or end-states (that which agents endeavor to bring about by the exercise of active power). If motives were taxonomic, however, then their influence would be causal; they would ‘push’ us to action in such a way that their causal characteristics could be stated in natural laws. But the influence of ends, Reid contends, is not causal; it is closer to that of advice or persuasion (EAP II.ii: 59). So, Reid concludes, motives are not taxonomic. Since moral features are motives, it follows that they are not taxonomic and, therefore, not natural.

Why does Reid think that the influence of motives is not causal? According to Reid, there are two ways that motives might make a causal contribution to action: they might be efficient (i.e., agent) causes, or occasional (i.e., nomic) causes of action. Reid thinks that neither possibility holds of motives. With respect to the first, Reid says ‘the influence of motives is of a very different nature from that of efficient causes. They are neither causes nor agents’ (EAP IV.iv: 214). That is, motives are not powers themselves nor could they be agents endowed with active power. With respect to the second, the influence of motives is not that of occasional causality: there are no laws—at least of the sort that are the subject matter of the natural sciences—that link motives to action. True, the system of necessity maintains that there are such laws, claiming ‘every action or change in action, in an intelligent being, is proportional to the force of motives impressed, and in the direction of that force’ (EAP IV.iv: 214). And Reid has a variety of critical things to say about attempts to identify analogues in the realm of action to Newton’s laws of motion, including that they tend to be vacuous (EAP IV.iv: 216–217). But his primary concern with such attempts is that in claiming that ‘motives are the causes, and the sole causes of actions... nothing is left to the agent but to be acted upon’ (EAP IV.iv: 217). To put the point in terms introduced earlier, Reid’s thought is that if nomic causality were sufficient to determine behavior, then there would be no further causal contribution needed by the agent. However, a theory that fails to specify properly the causal contribution of agents would fail to satisfy the two constraints on a satisfactory theory of action according to which agents must be able to make causal interventions of the right kind when acting.

The thrust of Reid’s thinking, then, is that any satisfactory account of the nature of moral features should focus on their role in action. Their role in action is that of an end: they are that in light of which agents pursue ends and by which agents regulate their behavior. These practical roles are, however, not causal in either of the senses that Reid identifies. When it comes to action, causal activity lies primarily

with the agent (and sometimes with the incitements that move agents) and not with the ends for, or in light of, which agents act. Of course, one could attempt to understand ends as playing a causal role in the production of action, but Reid's objection is that such efforts would render agents unable to make the sorts of causal interventions characteristic of genuine agency.

In laying out Reid's rationale for holding that moral features are not taxonomic and so not natural, our aim is (once again) not so much to defend what Reid says or to address gaps in his argumentation; it is instead to highlight the way in which Reid arrives at nonnaturalism. That way consists in treating considerations regarding the character of agency not as independent of, or falling out of, an account of the nature of moral features. Rather, the approach is thoroughly agency-centered in the sense that it asks, first, what agency is like and then addresses questions concerning what morality must be like given a proper understanding of agency.

5. Korsgaard's Kantian Alternative

We started our discussion by noting that Reid's agency-centered approach to ethical theorizing is, by the standards of both early modern and contemporary philosophy, highly unusual. Agency-centered strategies are primarily employed by Kantians, and these philosophers do not ordinarily identify themselves with the Reidian-Moorean tradition. To the contrary, they typically identify their views as both naturalistic and nonrealist (Walden [2012] and Korsgaard [n.d.] are explicit about this). We propose to close our discussion by considering how close these different types of agency-centered views are and where they ultimately diverge. (In this section we borrow from our discussion in Cuneo and Harp [2014].)

Comparing these different agency-centered views represents a shift in emphasis in our discussion, as we have primarily focused our attention on the nonnaturalist component of Reid's position to this point. Here, however, we focus on Reid's realist commitments and their bearing upon constitutivism. (These commitments have so far been partially apparent in the case that Reid offers for holding that self-regulation requires the existence of moral features.) In doing so, we hope for two things: first, to throw Reid's agency-centered approach into even sharper relief by bringing to the fore its constitutivist commitments and, second, to allay any lingering suspicions that the best (or only) way to support an agency-centered view is by embracing naturalistic, nonrealist metaethical views. Our strategy for accomplishing these aims is to employ Christine Korsgaard's nonrealist Kantian position as an illustrative comparison.

In the introduction to her *Self-Constitution*, Korsgaard (2009) writes that the name she gives to that feature 'which distinguishes human life from the lives of other animals' is the traditional one, namely, rationality:

As I understand it, reason is a power we have in virtue of a certain type of self-consciousness—consciousness of the grounds of our own beliefs and actions. This form of self-consciousness gives us a capacity to control and direct our beliefs and actions that the other animals

lack, and makes us active in a way that they are not.... This form of self-consciousness makes it necessary to *take control* of our beliefs and actions, but we must then work out how to do that: we must find normative principles, laws, to govern what we believe and do. The distinctive feature of human beings, reason, is therefore the capacity for normative self-government. (Korsgaard 2009: xi, emphasis in the original)

The similarities with some of the passages we have quoted from Reid (and indeed Reid's own introduction to the *Active Powers*) are striking; even the vocabulary of active agency, law, self-government are the same.

The similarities, moreover, do not stop here. Korsgaard presents a version of the disappearing agent objection, writing: 'to regard some movement of my mind or my body as *my action*, I must see it as an expression of my self as a whole, rather than as a product of some force that is at work *on* me or *in* me' (Korsgaard 2009: 18, emphasis in original). Like Reid, Korsgaard notes that instincts and incentives do not 'determine how we respond' to them; they '*propose* responses, but we may or may not act in the way they propose.... An incentive is experienced not as a force or a necessity but as a proposal' (Korsgaard 2009: 116, 119, emphasis original). When incentives do give rise to action, it is because of the 'intervention of principles' on the basis of which agents act (Korsgaard 2009: 121).

Of course, it would be nothing short of astonishing if the details of Korsgaard's project and Reid's view were substantially to align, given that Korsgaard gives no indication that she takes her view to be Reidian in any sense. And, sure enough, their views diverge in some important ways.

For one thing, Reid and Korsgaard offer different answers to what we earlier called the action problem. As Korsgaard sees things, the relationship that actions bear to agents is the reverse of that defended by Reid. Rather than explain action in terms of agents, Korsgaard endeavors to explain agency in terms of actions:

The intimate connection between person and action does not rest in the fact that action is caused by the most essential part of the person, but rather in the fact that the most essential part of the person is *constituted* by her actions. (Korsgaard 2009: 100, emphasis in original)

According to this view, something is an action, and indeed a particular agent's action, just in case that action is such as to constitute that agent.

Let us spell out this idea, as it is cryptic. Actions, says Korsgaard, (causally) connect agents to ends, because ends are that for which one acts. In addition, actions also constitute agents, where the 'task of self-constitution... involves finding some roles and fulfilling them with integrity and dedication. It also involves integrating those roles into a single identity, into a coherent life' (Korsgaard 2009: 25). Self-constitution, then, is the project of identifying and orienting one's life around projects and values of one's choosing. When an agent acts in such a way that she endorses these projects and values, she identifies with them, thereby constituting her practical identity or self.

For human beings, self-constitution occurs by way of conforming to standards of rationality. Like Reid, Korsgaard holds that there are two types of rational principles. However, she thinks of these principles rather differently from Reid, identifying them with the hypothetical and the categorical imperatives, respectively. The hypothetical imperative commands us to take the necessary means toward our ends; this corresponds to what Korsgaard calls the norm of efficacy because it is only by taking the necessary means that we can successfully bring about our ends. The categorical imperative, in contrast, corresponds to what Korsgaard calls the norm of autonomy: it governs the choice of actions by posing an admissibility test for acts being taken in pursuit of ends, evaluating both the act and the end together. Korsgaard also suggests that in satisfying the categorical imperative, we make ourselves both an agent and the cause of our desired end, where this causality is attributed to our own agency. The suggestion seems to be that when we satisfy the categorical imperative, we represent the principle that underlies our action as one that is rationally permissible for us *as an agent* to act on, which means that we are willing to have that principle constitute us (Korsgaard 2009: 131). An implication of this position is that agency is, in Korsgaard's view, episodic. When we violate the categorical imperative, as we often do, we fail to be agents.

We have enough of Korsgaard's view before us to see that there are some important differences between her view and Reid's. One difference, which can be concealed by Korsgaard's language of agency and her advocacy of the disappearing agent objection, is that Korsgaard does not accept anything like Reid's notion of active power. Instead, at points she seems to identify agency with the rational principles on the basis of which agents act, and sometimes she personifies them, as when she writes that incentives give rise to action because of the 'intervention of principles'. In Reid's view, this would be a mistake. Not only can the agent not be identified with principles, these principles no more apply themselves when an agent acts than intentions enact themselves when an agent acts; instead, agents must apply them. Differences such as this notwithstanding, the similarities between Reid's and Korsgaard's view are striking enough to raise the question why, given her account of agency and action, Korsgaard fails to draw Moorean conclusions similar to Reid's, instead opting for a version of naturalistic nonrealism whereby we impart values onto the world (see Korsgaard 1996).

At certain points, Korsgaard writes as if moral nonrealism simply falls out of her constitutivism according to which conforming to rational principles of certain kinds is constitutive of both action and agency. But at other places, Korsgaard articulates why she thinks positions such as Reid's cannot be correct. Targeting rationalists such as Samuel Clarke and Richard Price, Korsgaard argues that realism (which she takes to be equivalent to rationalism) cannot account for the bindingness or rational authority of practical principles such as the hypothetical imperative. It is worth quoting what she says at length:

The realist supposes that there are eternal normative verities of some sort—facts about which act-types or actions are right, or facts about what counts as a reason for what. How do we act on these verities? Apparently, by applying them in particular instances. We apply our

knowledge that an action is right by choosing it.... But notice that this sort of account could not possibly explain the normativity of the hypothetical imperative. We can see this by thinking about how it would have to work. The agent would have to recognize, as some sort of eternal normative verity, that it is good, or that it is required, or that there is a reason, to take the means to his ends.... How does he act on this recognition? How does he apply it to the case at hand?... We cannot explain how we are motivated to act on the hypothetical imperative, much less how we are bound by it, by appealing to the hypothetical imperative *itself*.... The point is that the hypothetical imperative cannot be a normative truth that we *apply* in practice, because it is the principle *in accordance with* which we are acting when we apply truths in practice. (Korsgaard 2009: 64–65, emphasis in original)

On this occasion, we propose to bracket discussion of the various claims and commitments that Korsgaard attributes to realism, focusing instead on the following issue.

Suppose Reid were to acknowledge that the hypothetical imperative is a fundamental norm of practical reasoning. Would he defend the authority of this principle in the way that Korsgaard claims he must, namely, by insisting that agents must recognize it as ‘some sort of eternal normative verity’? Not obviously. True, a realist such as Reid would hold that the hypothetical imperative is a mind-independent normative fact, as Korsgaard does not (although at no point does Reid identify a principle resembling the hypothetical imperative). But he, no less than the Kantian constitutivist, could claim that conforming to the hypothetical imperative is constitutive of practical reasoning. There is nothing about Reid’s nonnaturalist realist ontology that would suggest otherwise.

Not only could Reid claim this, in fact he claims something very much like this. Consider what Reid says about the principles of common sense. The principles of common sense, in Reid’s view, are those propositions that all well-formed, mature human persons must take for granted in their everyday lives on pain of suffering from serious normative defects, such as being incoherent, failing to treat like cases alike, and so forth. Regarding reasoning, for example, Reid writes: ‘A man who perfectly understood a just syllogism, without believing that the conclusion follows from the premises, would be a greater monster than a man born without hands or feet’ (EIP VI.v: x; cf. Ferrero 2009: 309, n.14). And regarding moral agency, Reid writes:

It is a first principle of morals, that we ought not to do to another, what we should think wrong to be done to us in like circumstances. If a man is not capable of perceiving this in his cool moments, when he reflects seriously, he is not a moral agent, nor is he capable of being convinced of it by reasoning.

From what topic can you reason with such a man? You may possibly convince him by reasoning, that it is his interest to observe this rule; but

this is not to convince him that it is his duty. To reason about justice with a man who sees nothing to be just or unjust; or about benevolence with a man who sees nothing in benevolence preferable to malice, is like reasoning with a blind man about colour, or with a deaf man about sound. (EAP III.iii.vi: 178)

And also:

If any man could say with sincerity, that he is conscious of no obligation to consult his own present and future happiness; to be faithful to his engagements, to obey his Maker, to injure no man; I know not what reasoning, either probable or demonstrative, I could use to convince him of any moral duty. As you cannot reason in mathematics with a man who denies the axioms, as little can you reason with a man in morals who denies the first principles of morals. The man who does not, by the light of his own mind, perceive some things in conduct to be right, and others to be wrong, is as incapable of reasoning about morals, as a blind man is about colours. Such a man, if any such man ever was, would be no moral agent. (EIP VII.ii: 723)

To be sure, there are some important differences between Reid's constitutivism as it is expressed in these passages and Korsgaard's. Reid does not claim that someone's failing to conform to the first principles of morals implies that she fails to be an agent altogether. Rather, he claims that when an agent fails to conform to these principles, she fails to be a *moral* agent. Moreover, it is true that Reid does not hold that 'only formal principles' such as the hypothetical imperative 'can be directly normative', as Korsgaard does, and that 'substantive principles must be derivable from formal ones if they are to be binding on the will' (Korsgaard 2009: 46). While these differences are significant, they are compatible with the position that, according to Reid's realism, some relatively formal normative principles are authoritative in the sense that Korsgaard wishes to defend and that, if an agent were to flout a sufficient number of them, she would cease to be an agent. It follows from this that both nonnaturalist realist constitutivists such as Reid and naturalist nonrealist constitutivists such as Korsgaard can (and do) accept claims about the constitutive status of normative principles. On this matter, there is no important difference between these views. Of course, this does not imply that Korsgaard has made a mistake in not embracing nonnaturalist realism, given her view of agency. But it does imply that her arguments for rejecting it do not hit their target. For what Reid's example establishes is that there are agency-centered approaches to ethical theorizing that yield positions that are at once nonnaturalist, realist, and constitutivist.

References

- Callergård, Robert. (2010) 'Thomas Reid's Newtonian Theism: His Differences with the Classical Arguments of Richard Bentley and William Whitson'. *Studies in History and Philosophy of Science*, 41, 109–19.
- Clarke, Samuel. (1998) *A Demonstration of the Being and Attributes of God and Other Writings*. Edited by Ezio Vailati. Cambridge, UK: Cambridge University Press.
- Cuneo, Terence. (2004) 'Reid's Moral Philosophy'. In Terence Cuneo and René van Woudenberg (eds.), *The Cambridge Companion to Thomas Reid* (Cambridge, UK: Cambridge University Press), 243–66.
- Cuneo, Terence, and Randall Harp. (2016) 'Thomas Reid'. In Kevin Timpe, Meghan Griffith, and Neil Levy (eds.), *The Routledge Companion to Free Will* (New York: Routledge), 332–42.
- Cuneo, Terence, and Randall Harp. (2014) 'Critical Review of Christine Korsgaard, *Self-Constitution*'. *Journal of Moral Philosophy*, 11, 97–110
- Darwall, Stephen. (1995) *The British Moralists and the Internal 'Ought'*. Cambridge, UK: Cambridge University Press.
- Enoch, David. (2011) *Taking Morality Seriously*. Oxford: Oxford University Press.
- Ferrero, Luca. (2009) 'Constitutivism and the Inescapability of Agency'. In Russ Shafer-Landau (ed.), *Oxford Studies in Metaethics*, 4, 304–33.
- Fodor, Jerry. (1991) 'A Modal Argument for Narrow Content'. *Journal of Philosophy*, 88, 5–26.
- Gill, Michael. (2006) *The British Moralists on Human Nature and the Birth of Secular Ethics*. Cambridge, UK: Cambridge University Press.
- Harris, James. (2005) *Of Liberty and Necessity*. Oxford: Oxford University Press.
- Katsafanas, Paul. (2013) *Agency and the Foundations of Ethics: Nietzschean Constitutivism*. Oxford: Oxford University Press.
- Korsgaard, Christine. (n.d.) 'Animal Selves and the Good'.
- Korsgaard, Christine. (2009) *Self-Constitution*. Oxford: Oxford University Press.
- Korsgaard, Christine. (1996) *The Sources of Normativity*. Cambridge, UK: Cambridge University Press.
- Moore, G. E. (1903/1993) *Principia Ethica*. Cambridge, UK: Cambridge University Press.
- McNaughton, David. (2013) 'Butler's Ethics'. In Roger Crisp (ed.), *The Oxford Handbook to the History of Ethics* (Oxford: Oxford University Press), 377–98.
- Price, Richard. (1787/1948) *A Review of the Principal Questions in Morals*. Edited by D. D. Raphael. Oxford: Oxford University Press.
- Reid, Thomas. (1788/2010) *Essays on the Active Powers of the Human Mind*. Edited by Knud Haakonssen and James Harris. Edinburgh: Edinburgh University Press (abbreviated as EAP).
- Reid, Thomas. (1785/2002) *Essays on the Intellectual Powers of Man*. Edited by Derek R. Brookes. Edinburgh: Edinburgh University Press (abbreviated as EIP).
- Reid, Thomas. (2002) *The Correspondence of Thomas Reid*. Edited by Paul, Wood. Edinburgh: Edinburgh University Press.
- Reid, Thomas. (1995) *Thomas Reid on the Animate Creation: Papers Relating to the Life Sciences*. Edited by Paul, Wood. Edinburgh: Edinburgh University Press.
- Velleman, J. D. (1992) 'What Happens When Someone Acts?' *Mind*, 101, 461–81.
- Velleman, J. D. (2009) *How We Get Along*. Cambridge, UK: Cambridge University Press.
- Walden, Kenny. (2012) 'Laws of Nature, Laws of Freedom, and the Social Construction of Normativity'. In Russ Shafer-Landau (ed.), *Oxford Studies in Metaethics*, 7, 37–79.
- Wiggins, David. (1993) 'A Neglected Position?' In John Haldane and Crispin Wright (eds.), *Reality, Representation and Projection* (Oxford: Oxford University Press), 329–38.
- Wood, Paul. (2004) 'Thomas Reid and the Culture of Science'. In Terence Cuneo and René van Woudenberg (eds.), *The Cambridge Companion to Reid* (Cambridge, UK: Cambridge University Press), 53–76.
- Yaffe, Gideon, (2004) *Manifest Activity: Thomas Reid's Theory of Action*. Oxford: Oxford University Press.