Forms of Rational Agency

DOUGLAS LAVIN

Abstract
A measure of good and bad is *internal* to something falling under it when that thing falls under the measure in virtue of *what it is*. The concept of an internal standard has broad application. Compare the external breed standards arbitrarily imposed at a dog show with the internal standards of health at work in the veterinarian's office. This paper is about practical standards, measures of acting well and badly, and so measures deployed in deliberation and choice. More specifically, it is about the attempt to explain the unconditional validity of certain norms (say, of justice and prudence) by showing them to be internal to our agency and the causality it involves. This is *constitutivism*. Its most prominent incarnations share a set of assumptions about the nature of agency and our knowledge of it: *conceptualism*, *formalism* and *absolutism*. This essay investigates the merits and viability of rejecting *all* of them while still seeking the ground of practical normativity in what we are, in our fundamental activity.

1. Introduction: reason and will

We all know Hume says,

> the impulse [to pursue a certain object] arises not from reason, but is only directed by it… Reason is, and ought only to be, the slave of the passions[1]

whereas Kant says,

> the true vocation of reason must be to produce a will that is good, not simply as a means to other purposes, but *in itself*.[2]

Thanks especially to Matthew Boyle. I am indebted as well to Matthias Hasse, Christine Korsgaard, Sergio Tenenbaum and Michael Thompson. I have benefited from discussion at the Institute of Philosophy, Leeds, Leipzig, LMU, and the DFG-Netzwerk on Praktisches Denken und gutes Handeln. I am grateful to the Alexander von Humboldt-Stiftung for generous financial support.

[1] David Hume, *A Treatise of Human Nature*, ed. D. F. Norton and M. J. Norton (Oxford: Oxford University Press, 2003), II.3.iii.

[2] Immanuel Kant, *Groundwork of the Metaphysics of Morals*, ed. and trans. Mary Gregor (Cambridge: Cambridge University Press, 1998), 4:396.

doi:10.1017/S1358246117000091 ©The Royal Institute of Philosophy and the contributors 2017

https://doi.org/10.1017/S1358246117000091 Published online by Cambridge University Press

They are commonly taken to be in a dispute about the role of reason in practice: whether this consists merely in determining means suitable to ends given from elsewhere, or also, and more fundamentally, in determining what ends to be pursued. How do we adjudicate such a dispute? When one party claims that rational reflection can and should play only an instrumental role in deliberation, while another claims that it can and should determine what ends are to be pursued, where do we look to figure out who is right?

An influential body of work in contemporary ethics and practical reason proposes that the dispute can be resolved by reflecting the nature of agency. The idea is that we might establish the authority of certain practical requirements – perhaps only conditional requirements between ends and means, or perhaps unconditional requirements on our ends themselves – by deriving them from the the very idea of agency. Seen through this lens, the dispute between Hume and Kant is about whether an agent, considered simply *as such*, is subject only to requirements to take means conducive to its ends, or whether it must also be subject to further, substantive requirements on the ends it pursues. The dispute, in other words, concerns what standards of correctness in practical thought, or measures of acting well, belong *internally* or *constitutively* to the power to act.

On one widespread conception, *constitutivism* is the project of defending a thesis of precisely this sort: a thesis that basic practical requirements are grounded in the nature of action.[3] Constitutivists typically argue, in addition, that fundamental practical principles can *only* be validated in this way – that an agent could shrug off without inconsistency any requirement that could not be traced back to the very nature of agency. Even in the absence of such an argument, this way of understanding the authority of a standard of action has obvious appeal. If it could be shown that any agent is, by nature, subject to certain requirements, this would elegantly clarify why we, who know ourselves to act intentionally, must (in some sense) know ourselves to be bound by them. And if there were such a proof, it would provide a particularly attractive sort of vindication: it would show them to be objective, while at the same

[3] Thus in an article summarising the current discussion, Elijah Millgram writes '[c]onstitutivist arguments move from the premises that anything you *do* will inevitably be an action, and that an action is such-and-such, to the conclusions that whatever you *do* will be a such-and-such' Elijah Millgram, 'Pluralism About Action', in *A Companion to the Philosophy of Action*, ed. Timothy O'Connor and Constantine Sandis (Oxford: Wiley-Blackwell, 2010), 90.

https://doi.org/10.1017/S1358246117000091 Published online by Cambridge University Press

time illuminating why they are such as to immediately engage the will of each (perhaps not always, but when we deliberate without interference). It would thus avoid a familiar dilemma: the apparent need to choose between views which vindicate the objectivity of standards of action in a way that leaves a residual mystery about how they might make a dent on the will, and views which explain the immediate bearing of normative thought to move us to act on motivation, but in a way that deprives them of objectivity. Even if some philosophers have given up one or the other, they have done so under the pressure of the difficulty. The natural first response is to seek an account that shows the incompatibility to be merely apparent.

Can the concept of action bear the theoretical weight constitutivists place on it? One ground for skepticism is that different constitutivists come to such strikingly different conclusions about which requirements are built into the structure of agency. Humeans insist that only hypothetical requirements to take appropriate means to one's actual ends, whatever they may be, can be part of this structure.[4] Kantians maintain, on the other hand, that being subject to categorical requirements on ends is a condition of the possibility of being an agent at all.[5] There is reason to worry that intuitions about the requirements contained in the concept of agency simply reflect substantive convictions about the normative requirements to which we are subject, and cannot supply an independent point of leverage on the latter question. Moreover, even if it were shown that, on some

[4] James Dreier, 'Humean Doubts About the Practical Justification of Morality', in *Ethics and Practical Reason*, ed. Garrett Cullity and Gaut Berys (Oxford: Clarendon Press, 1997), 81–100; Sharon Street, 'What Is Constructivism in Ethics and Metaethics?' *Philosophy Compass* **5**(5) (2010): 363–84; Sharon Street, 'Coming to Terms with Contingency: Humean Constructivism About Practical Reason', in *Constructivism in Practical Philosophy*, ed. Jimmy Lenman and Yonatan Shemmer (Oxford University Press, 2012); Candace A. Vogler, *Reasonably Vicious* (Harvard University Press, 2002).

[5] Christine M. Korsgaard, *Self-Constitution: Agency, Identity, and Integrity* (Oxford University Press, 2009); Christine M. Korsgaard, *The Constitution of Agency: Essays on Practical Reason and Moral Psychology* (Oxford University Press, 2008); Sebastian Rödl, *Self-Consciousness* (Harvard University Press, 2007). Michael Smith, 'The Magic of Constitutivism', *American Philosophical Quarterly* **52**(2) (2015); Michael Smith, 'A Constitutivist Theory of Reasons: Its Promise and Parts', *Law, Ethics and Philosophy* **1**(0) (2013): 9–30; David J. Velleman, *The Possibility of Practical Reason* (Oxford University Press, 2000), and *How We Get Along* (Cambridge University Press, 2009).

https://doi.org/10.1017/S1358246117000091 Published online by Cambridge University Press

conception of agency, being an agent involves being bound by certain principles, it might still be asked *what reason there is to be an agent in the relevant sense*. A number of critics have argued that, in the absence of an answer, a putative derivation of normative requirements from the concept of agency cannot show them to be *unconditionally* binding. And, the critics add, answering this question requires appealing to a source of normativity other than, or external to, agency itself, and so the constitutivist program simply cannot succeed.[6]

2. Three assumptions about constitutivism

The aim of the present essay is to raise some questions about a set of assumptions accepted by all the parties to this dispute – Kantians, Humeans and their rationalist critics. They assume that an adequate constitutivist account must meet the following conditions:

> **Conceptualism:** It must show certain requirements on action to be *analytically contained* in the very concept of agency, a concept that applies to all subjects capable of acting for reasons, however different they may otherwise be.
>
> **Formalism:** It must show how to *derive* the relevant substantive requirements from a conception of agency, in such a way that a subject who was skeptical of them, but who understood himself to be an agent in the relevant sense, and was capable of appreciating what this implies, would be compelled to accept the requirements on pain of inconsistency.
>
> **Absolutism:** It must show that certain normative requirement hold for *all possible* agents-capable-of-acting-for-reasons. If a kind of rational agent which is not subject to these requirements were shown to be possible, the relevant constitutivist project would be shown to be a failure.

These assumptions are closely related. Someone who accepts Conceptualism must accept both Formalism and Absolutism, and while a commitment to Formalism or Absolutism may not strictly imply the others, it certainly makes them natural to think. In any

[6] David Enoch, 'Agency, Shmagency: Why Normativity Won't Come from What Is Constitutive of Action', *The Philosophical Review* **115**(2) (2006): 169–98; Elijah Millgram, 'Practical Reason and the Structure of Actions', *Stanford Encyclopedia of Philosophy*, 2008; Peter Railton, 'On the Hypothetical and Non-Hypothetical in Reasoning About Belief and Action', in *Ethics and Practical Reason*, 53–79; T. M. Scanlon, 'The Appeal and Limits of Constructivism', in *Constructivism in Practical Philosophy*; R. Jay Wallace, 'Constructivism About Normativity: Some Pitfalls', in *Constructivism in Practical Philosophy*.

https://doi.org/10.1017/S1358246117000091 Published online by Cambridge University Press

case, my project here is not to address subtle issues about the logical relations between these commitments, but to ask whether *any* should be accepted.

I raise this question from a standpoint deeply sympathetic to the spirit of constitutivism. I think it is right to see the debate between Hume and Kant as a debate about the kinds of practical requirement to which we are subject in virtue of our capacity to act for reasons. And I agree that understanding the basis of practical normativity in this way would provide a particularly elegant and attractive way of addressing various basic problems in ethics and meta-ethics. My aim is to show that the project of constitutivism can be understood more broadly than its contemporary advocates and opponents typically assume, a way that retains its deepest attractions, while avoiding various difficulties that beset its most prominent contemporary incarnations. If I am right, there is room for a pluralist, non-formalist constitutivism that that does not seek to ground normative requirements on action in the sheer concept of rational agency, but that still presents appealing and distinctive responses to fundamental questions in ethical theory. Moreover, seeing this possibility will shed a revealing light on the debate between Hume and Kant, and the kind of resolution it might admit.

My conviction that such a view is possible grows out of reflection on Aristotelian ethics. Contemporary discussions of constitutivism seldom mention Aristotle and the kind of naturalism he espouses, but he surely deserves to count as a constitutivist in some important sense.[7] His view is, famously, that the fundamental norms governing how to lead our lives are grounded in the fact that we have by nature a certain *ergon* or characterisitic activity, and that the structure of the human soul implies a certain appropriate division of labor in how do what it is our to do. Thus, for Aristotle, the notion of the *nature* of a human being plays a crucial role in explaining practical normativity, or again the good in action: the fact that we, human beings, act well in doing in certain ways is grounded, ultimately, in what we essentially are, where this is characterised in terms of what we

[7] I do not mean to say Aristotle goes entirely unmentioned. See, for example, Stephen Engstrom, 'The Complete Object of Practical Knowledge', in *The Highest Good in Aristotle and Kant*, ed. Joachim Aufderheide and Ralf M. Bader (Oxford University Press, 2015), 129–57; Christine M. Korsgaard, 'Aristotle's Function Argument', in *The Constitution of Agency* and 'How to be an Aristotelian Kantian Constitutivist' (unpublished).

https://doi.org/10.1017/S1358246117000091 Published online by Cambridge University Press

essentially *do*, our characteristic *activity*.[8] In this sense, Aristotle holds that the norms to which we are subject are *constitutive* norms. It seems, however, that his conception of our constitution, of what we are, includes more than the mere fact that we are capable of acting for reasons: he locates the ground of the practical norms to which we are subject, not in our nature as agents, but in our nature as human beings, a certain kind of material being. Moreover, he displays a striking lack of concern with the task of showing that the relevant norms apply to all possible agents, still less with the task of deriving the relevant norms from some abstract or formal characterization of our nature. He famously says that he aims to speak only to persons who have received a sort of upbringing that already equips them to see the point of certain ways of acting, and he seems to allow for the possibility that different life-forms might imply different lists of ethical virtues. I think these features of Aristotle's project are not simply independent commitments, but expressions of a single underlying idea – *hylomorphism*, to give it a name. A central aim in what follows is to bring out the bearing of this idea on the contemporary debate about constitutivism.

3. Absolutist constitutivism: Kantians and anti-Kantians

3.1. Kantian absolutism

It is a brute but inescapable fact that human beings confront the question how to act. We can deliberate about what speaks in favor of acting one way or another another, but the question whether to be agents at all seems not to be one we can coherently regard as open to deliberation. Christine Korsgaard puts the point this way:

> Human beings are *condemned* to choice and action. Maybe you think you can avoid it, by resolutely standing still, refusing to act, refusing to move. But it's no use, for that will be something you have chosen to do, and then you will have acted after all… You have no choice but to choose, and to act on your choice.[9]

Absolutist constitutivists seek to ground practical normativity on this fixed point. If it could be shown that certain normative requirements on action are implied by the very exercise of agency, then, they argue, these requirements would be shown to be inescapable in a special and

[8] Aryeh Kosman, *The Activity of Being: An Essay on Aristotle's Ontology* (Harvard, 2013).
[9] Korsgaard, *Self-Constitution*, 1.

https://doi.org/10.1017/S1358246117000091 Published online by Cambridge University Press

interesting sense. For we would be committed to them, not in a manner conditional on our having one or another end, but unconditionally, insofar as we act in pursuit of any end whatsoever. Normative requirements grounded in the nature of agency would thus be unconditional requirements on action.

Kant's *Groundwork* is a principal source of inspiration for absolutist constitutivism. On an influential interpretation, Kant there seeks to prove that a certain fundamental norm governing action, the categorical imperative, is implicit in the very idea of a will, conceived a capacity to determine oneself to act on the basis of reasons. *Kantian* constitutivists seek to defend some version of this claim, and thus to vindicate the idea that there are requirements to which we are subject unconditionally, simply as rational agents. *Anti-Kantian* constitutivists deny that any unconditional practical requirements can be wrung from the sheer idea of rational agency. Many stop short of Hume's conclusion that there can be no rational basis for assessing our ultimate ends, but they maintain, in any case, that the very idea of a capacity to act for reasons cannot be the source of the norms governing such an assessment. Surveying this dispute will help to clarify the attractions of absolutist constitutivism, and also to bring out some questionable assumptions on which it rests.

Korsgaard is the leading contemporary advocate of the Kantian position. In a series of influential papers and monographs, she has sought to show that Kant's categorical imperative articulates a necessary norm of agency – a norm to which rational agents are committed by their very exercise of agency, so that in failing to conform to it, they have acted badly by their own lights. For present purposes, the aspect of Korsgaard's argument of interest is not her case for Kant's categorical imperative in particular, but her general strategy for showing that some such unconditional normative requirement holds.

The strategy, very roughly, is to argue that, unless some unconditional normative requirement holds, there can be no distinction between an exercise of agency, attributable to the agent herself, and the mere expression of forces operative in a person, producing behavior that does not constitute the agent's determining herself to do something.

To exercise agency, Korsgaard proposes, is to determine oneself to perform some act A for the sake of some end E. (She reserves the term 'action' for this form of endeavor – doing A for E – and uses the term 'act' to refer to the thing done – do A). But if an agent is to exercise agency in this sense, she argues, there must be a basis for distinguishing the agent's determining to do A for E from some mere impulse determining her to do A for E. If some impulse in the agent, say an irresistible urge for E, determines her to do A for E, then *she* does

https://doi.org/10.1017/S1358246117000091 Published online by Cambridge University Press

not determine her action; rather, she is determined to act in a certain way by a force whose efficacy does not depend on her will. There is, however, a basis for distinguishing the agent's determining her own action from her being non-agentially determined to act only if the agent can consider the question whether to do *A* for *E*. And this consideration, Korsgaard goes on to argue, must take the form of assessing whether this 'maxim' – do *A* for *E* – can be willed as a universal law. So, in effect, to be capable of exercising agency at all, an agent must be capable of assessing whether her maxims satisfy the categorical imperative, and in representing herself as choosing to perform a certain action, she represents herself as taking her maxim to meet this test.

This line of reasoning would be interesting, I think, even if its final steps were judged unsuccessful. Whatever we think of Korsgaard's concluding argument (which I have not tried to reconstruct) for the claim that agents are always already assessing their actions by the categorical imperative, I think we should be intrigued by her strategy for showing that agency requires, not just the capacity to take reasonable means to one's ends, but the capacity to make some sort of unconditional rational assessment of one's action as a whole (one's doing *A* for *E*). Korsgaard takes this point by itself to rule out a Humean conception of practical reason. Humeans hold that the proper function of practical reason is simply to consider what means should be taken to realize given ends. But according to Korsgaard, 'the instrumental principle (i.e., the principle that one should take the necessary means to one's ends) cannot stand alone': if there were no supporting rational commitment to pursue some end *E*, there would be no basis for a distinction between the agent's failing to conform to the instrumental principle (by not taking the necessary means to achieve *E*) and her simply ceasing to pursue *E*.[10] For a principle to be a genuine norm governing action, Korsgaard holds, it must be possible for an agent to fail to conform to this principle. In particular, if the instrumental principle is to be genuinely normative, it must be possible for an agent who is pursuing *E* to know that doing *A* is necessary for achieving *E*, and yet fail to do *A*. But if the agent is not rationally committed to pursuing *E*, then her not taking what she knows to be the necessary means to *E* does not exhibit any rational failure, for nothing about the agent's situation requires that she should pursue *E* at all. Korsgaard concludes that, in the absence of some sort of rational commitment to the pursuit of particular ends, there is no such thing as the agent's being instrumentally irrational.

[10] Christine M. Korsgaard, 'The Normativity of Instrumental Reason', in *Ethics and Practical Reason*, 251.

https://doi.org/10.1017/S1358246117000091 Published online by Cambridge University Press

This result, she maintains, is just another manifestation of the point that the applicability of the concept of agency requires a distinction between the agent's determining herself to action and something else determining this:

> [Hume] has no resources for distinguishing a person's ends from what she actually pursues. Another way to put the same point… is to say that Hume has no resources for distinguishing the activity of the person *herself* from the operation of beliefs, desires, and other forces *in her*.[11]

I'm describing these arguments because I think they have undeniable appeal. There is something attractive in the thought that agency requires determining oneself to act in pursuit of a certain end, as opposed to being passively determined to action; and it seems plausible that, if we could not regard certain ends as rational to pursue, the question whether it is rational to take certain means would not have the sort of significance it has for us. Nevertheless, I believe that each of these ideas is in fact question-begging in the context of the debate with anti-Kantians. Explaining why this is so will bring out the need for an alternative account of the appeal of Korsgaard's arguments, one that does not require interpreting them as stating what must be so where there is action at all.

Consider the first idea: that agency requires determining oneself to act in pursuit of a certain end. In arguing from this conception of agency to the existence of unconditional requirements, Korsgaard presupposes that *what* an agent must determine herself to do is of the form: do A for E.[12] This is to conceive of the object of choice, not simply as an act (do A), but a certain principle of action (do A for E). But this is precisely what anti-Kantians should not admit. On their view, the only necessary requirement on agency is the conditional requirement that, given a certain end, the agent should take appropriate means. Hence they should not admit that agents, simply as such, face rationally-assessable choices about whether to act on a certain principle; for this is tantamount to admitting that agents, simply as such, face rationally-assessable choices about whether to pursue certain ends. Anti-Kantians should admit only that, *given the ends* an agent is in fact going after, she faces the choice *whether to take certain means*. On this view, the agent's end does not fall

[11] Ibid., 233.

[12] Thus Korsgaard takes it to be uncontentious that 'what the will chooses is, strictly speaking, actions', where an 'action' is a doing of A for E. That this assumption is contentious has been discussed by Millgram in 'Pluralism about Action', though he interprets its significance very differently than I do.

https://doi.org/10.1017/S1358246117000091 Published online by Cambridge University Press

within the scope of her choice – or at any rate, the sheer fact that she is an agent does not entail that it do so. And so, if Korsgaard's argument depends on assuming that the object of choice is what *she* calls 'action' – doing A for E – then it takes for granted the very point in dispute. But once this assumption is dropped, there is no direct route to the conclusion that an agent whose ends simply appear and disappear as the result of forces operative in her does not *determine herself* to act. After all, on the anti-Kantian view, the requirement of self-determination applies, not to the agent's principle of *doing A for E*, but only to her act of doing A, and no reason has been given to think that the agent's doing A (rather than, say, B, which she judges to be possible but less conducive to achieving E) cannot be self-determined.

Turn now to the second idea: that, if an agent could not regard certain ends as rational to pursue, the question whether it is rational to take necessary means would lose significance for her. Korsgaard's argument for this presupposes a certain conception of the role that the instrumental principle must play in our practical thought. As we have seen, she assumes that it must be possible for an agent to be pursuing E, and know (or at least believe) that doing A is necessary for E, and yet not do A. Only this, she thinks, would be a genuine failure to respect the instrumental principle, and only a principle that an agent can fail to respect can be a genuine norm of agency. But again, this is precisely what anti-Kantians should not admit. On their view, an agent, considered simply as such, is subject only to the norm that,

($\mathrm{IP}_{\mathrm{AntiK}}$) Given the ends she is pursuing, she should: take the necessary means.

Here the 'should' has narrow scope: it applies only to her taking the means that are in fact necessary to achieve her ends.[13] Contrast this with a wide-scope reading of the instrumental principle:

[13] The scope contrast I am drawing here should not be confused with the more familiar contrast between a wide scope reading of practical norms on which they are of the form:

(WS) It ought to be the case that, if p, then the agent does A.

and a narrow scope reading on which their form is:

(NS) If p, then it ought to be the case that the agent does A.

Korsgaard holds that fundamental norms of practical rationality cannot be formulated in the manner of (WS), on pain of their not being genuinely practical norms ('The Activity of Reason', *Proceedings and Addresses of the American Philosophical Association* **83**(2) (2009): 27–29.) It is less clear whether she would accept formulations along the lines of (NS) or hold

https://doi.org/10.1017/S1358246117000091 Published online by Cambridge University Press

(IP_K) The agent should: take the necessary means to her ends.

Let it be granted that a principle can only count as a genuine norm of agency if it is possible for an agent to violate it. For an agent to violate (IP_K), she would need to choose *not to take the necessary means to her ends*; and this plausibly requires her believing of some act A that it is necessary for achieving her end E and yet not doing A. But for an agent to violate (IP_{AntiK}), she would need only to be pursuing a certain end E and choose *not to do* A, which is in fact necessary for E. To impose the additional requirement that she should do this while believing that doing A is necessary to achieve E is to require, not only that she should be capable of voluntarily acting in a manner that is (in fact) not conducive to her end, but that she should be capable of voluntarily flouting the instrumental principle itself – knowing it to apply and yet choosing not to conform to it. But this is again tantamount to assuming the very point in dispute. Anti-Kantians simply should not admit that rational agency requires being 'governed by the instrumental principle' in a sense that would require the possibility of *this* specific sort of violation. For on their view, the only necessary object of practical reason is: whether *to do such-and-such*, not whether *to conform to the principle that requires doing such-and-such*.

The upshot of these criticisms is that Korsgaard has no non-question-begging argument for the conclusion that any agent, simply as an agent, is subject to some unconditional normative requirement. Her arguments apply, at best, only to a certain *form* of agent, one capable of choosing to do A for E – choosing to act on a certain principle, as I have put it. How serious a setback is this for Korsgaard's Kantian constitutivism? It is a decisive defeat *if* the Kantian project must be to derive unconditional normative requirements from the very idea of agency; but is it obvious that Korsgaard's arguments are of interest only if they meet this condition? Whatever may be true of agents simply as such, it seems attractive to hold that human agents are characteristically capable of choosing to act on principles. At any rate, I take myself to be capable of making such

---

out for something else altogether. But whatever the outcome of this dispute, there remains the question whether the scope of the instrumental principle is wide or narrow in the sense I have distinguished: whether it says that we are rationally required to take the means to our ends, or only that, given certain ends, we are rationally required to take certain means. My claim is that, with respect to this latter contrast, Korsgaard's argument presupposes the wide reading.

https://doi.org/10.1017/S1358246117000091 Published online by Cambridge University Press

choices – and I take myself to have this capacity, not on the basis of self-observation, but by reflection on my very ability to understand and deliberate about whether it is right to act on the principle: do *A* for *E*. I don't claim any special success in subjecting my principles of action to such scrutiny: I do not often rise to this level of reflection, nor do I claim any distinction in disciplining my action by its results. But I know myself at least to be *capable* of this sort of self-scrutiny: embarrassment over doing so only here and there expresses a presumption that I could do better.[14] And to this extent, it seems that Korsgaard's reflections – to the extent that they are cogent – might after all get a grip on me even if they do not get a grip on all conceivable agents.

3.2 Kant's anti-Kantianism

Contemporary Kantians like Korsgaard commonly suppose their burden is to show that any conceivable rational agent must accept the categorical imperative as a norm governing her action. One indication that this might not be necessary to comprehend vindication of comes from the historical Kant. For whatever his view might have been at the time of writing the Groundwork, by the time he wrote the *Critique of Practical Reason*, he plainly held that being subject to the moral law is known to us, not as a consequence of some more basic premise about the general nature of rational agency, but as a basic fact of reason – something known to us in virtue of having the sort of power of practical reason that we actually have, together with the capacity to reflect self-consciously on exercises of this power. Indeed, in the *Religion*, Kant says explicitly that the concept of moral 'personality' (i.e. the concept of 'a rational and at the same time responsible being') is not contained in the concept of

[14] Compare Kant's famous remark on how our very recognition of the moral law implies a recognition of our own practical capacities: 'Ask [someone] whether, if his prince demanded, on the threat of the… prompt penalty of death, that he give false testimony against an honest man whom the prince would like to ruin under specious pretenses, he might consider it possible to overcome his love of life, however great it may be. He will perhaps not venture to assure us whether or not he would overcome that love, be he must concede without hesitation that doing so would be possible for him. He judges, therefore, that he can do something because he is conscious that he ought to do it'. Immanuel Kant, *Critique of Practical Reason*, trans. Mary J. Gregor, (Cambridge University Press, 1997), 5:30.

https://doi.org/10.1017/S1358246117000091 Published online by Cambridge University Press

'humanity' (i.e., the concept of 'a living and at the same time rational being'):

> [F]rom the fact that a being has reason, it does not at all follow that, simply by virtue of representing its maxims as suited to universal legislation, this reason contains a faculty of determining the power of choice unconditionally, and hence to be 'practical' on its own; at least, not so far as we can see. The most rational being of this world might still need certain incentives, coming to him from the objects of inclination, to determine his power of choice. He might apply the most rational reflection to these objects – about what concerns their greatest sum as well as the means for attaining the goal determined through them – without thereby even suspecting the possibility of such a thing as the absolutely imperative moral law which announces to be itself an incentive, and, indeed, the highest incentive. Were this law not given to us from within, no amount of subtle reasoning on our part would produce it or win our power of choice over to it.[15]

Kant here appears to admit quite unequivocally that a merely Humean power of practical reason (i.e. one that can and should only reason instrumentally, and perhaps prudentially, on the basis of given sensible inclinations) is conceptually possible, or in any case that we cannot rule it out. If so, the existence of unconditional practical norms cannot be demonstrated by any analytical-deductive argument beginning simply from the idea of a living being with the power of practical reason. Nevertheless, Kant holds that we actual humans know ourselves to be, not merely Humean agents, but morally responsible persons, on the basis of a law 'given to us from within' – namely, the moral law implicit in our very understanding of the unconditional practical 'ought'.[16]

Kant's position suggests the possibility of a non-absolutist constitutivism: one that recognizes the possibility of more than one form of agency, and that seeks to vindicate our subjection to unconditional normative requirements by showing them to belong constitutively

[15] Immanuel Kant, *Religion Within the Boundaries of Mere Reason and Other Writings*, trans. George Di Giovanni (Cambridge University Press, 1998), 6:26.

[16] Op. cit. note 145:31. This is commonly supposed to reflect a change in Kant's position. For doubts see Stephen Engstrom, *The Form of Practical Knowledge* (Harvard University Press, 2009). Sergio Tenenbaum, 'Speculative Mistakes and Ordinary Temptations: Kant on Instrumentalist Conceptions of Practical Reason', *History of Philosophy Quarterly* **20**(2) (2003), 203–23 and 'The Idea of Freedom and Moral Cognition in Groundwork III', *Philosophy and Phenomenological Research* **84**(3) (2012), 555–89.

https://doi.org/10.1017/S1358246117000091 Published online by Cambridge University Press

to our specific form of agency, rather than to rational agency *simpliciter*.

3.3. Humean absolutism

The possibility of such a pluralist constitutivism is routinely overlooked, not only by contemporary Kantians aiming to defend unconditional normative requirements, but also by their anti-Kantian opponents. To take just one example, consider[17] which begins by arguing that

1. Any rational agent, simply as such, is subject to the instrumental principle.

For, he argues, unless an agent's practical reasoning is governed by this principle, she is incapable of drawing any practical inferences from given aims to determinate actions, and if she is incapable of drawing practical inferences, then nothing can count as a practical reason for her. He then argues that

2. The instrumental principle is the only necessary principle of practical reason.

For, he holds, no other putative principle of practical rationality shares the unquestionable necessity of the instrumental principle: it makes no sense to ask for a reason to take the means to your ends, whereas this question at least makes sense for any other putative practical requirement. But it is plausible that

3. If compliance with morality were demanded by practical reason, this would require the existence of an unconditional, non-instrumental principle of practical reason.

Thus, Dreier concludes, Hume is vindicated:

4. Practical reason does not demand compliance with morality.

Without disputing any premise of this argument, I want to note an ambiguity in its use of the phrase 'practical reason'. It might mean either practical reason, considered simply as such or practical reason as we know it. The considerations Dreier offers in favor of premise (2) – granting their cogency for the sake of argument – establish this premise only if 'practical reason' is read in the former way: they establish that no further principle is required by the very idea of a capacity to reason practically. But the conclusion, (4), is a

[17] Dreier, 'Humean Doubts About the Practical Justification of Morality'.

https://doi.org/10.1017/S1358246117000091 Published online by Cambridge University Press

vindication of Hume only if 'practical reason' is read in the latter way, as a claim about our actual power of practical reason – for surely Hume meant to claim, not just that a Humean agent is conceptually possible, but that we are actually Humean agents, whose reason is and ought to be a slave to our passions.

4. The possibility of pluralism

To assume that a vindication of Hume can be inferred from Dreier's premises is, in effect, to assume that if only the instrumental principle is necessitated by the very concept of a capacity to reason practically, then the instrumental principle is the only principle that can belong necessarily to a capacity for practical reason. It is to assume that if no further requirement of practical rationality is conceptually necessary, then no further requirement of practical rationality is actually possible. This is to overlook the possibility of a *pluralist* constitutivism, one which acknowledges that a Humean form of practical reason is conceptually possible, and thus that no unconditional practical requirements can be derived from an analysis of the very concept of a rational agent, but which maintains that other, richer forms of agency are also possible.

Dreier does not give an argument to rule out this possibility: he simply assumes that the debate over constitutivism is a debate about whether there are practical requirements to which all conceivable rational agents are subject. And though I cannot argue the point here, I think this holds true also of other prominent contemporary anti-Kantians: they pass from arguments for the claim that Humean rational agents are conceptually possible to conclusions about what practical reason demands of us, by way of an uncritical assumption that any constitutivist vindication of practical requirements must ground those requirements in the very idea of a rational agent.[18] This absolutist assumption, shared by contemporary Kantians and anti-Kantians, gives their dispute the appearance of a conflict in which only one side can be correct. But once we recognize the possibility of a pluralist constitutivism, we can see that each side might be right about something significant. The anti-Kantians might be right that the very concept of a capacity to act for reasons does not imply any further requirement beyond the instrumental principle, while the Kantians might be right that our actual capacity to act for

[18] See, for example, Smith, 'The Magic of Constitutivism', Street, 'Coming to Terms with Contingency', Vogler, *Reasonably Vicious*.

https://doi.org/10.1017/S1358246117000091 Published online by Cambridge University Press

reasons takes a form that implies the existence of unconditional practical norms. So long as neither side frames its point in absolutist terms, their views are consistent.

5. The problem of inescapability

If the discussion of the last section is sound, a pluralist position – one that allows for the possibility of multiple forms of agency, each constitutively subject to its own characteristic requirements of practical rationality – is a dialectically attractive option that is mostly overlooked in contemporary debates over constitutivism. This calls for explanation: if this sort of position is available and attractive, why is it not considered?

To see a reason for this neglect, it will help to consider an influential criticism of the constitutivist program. This criticism brings out that, on some ways of understanding the thesis that being subject to certain norms is 'constitutive of being an agent', this thesis proves too flimsy to support the metaethical results that constitutivists hope to derive from it. It can seem that the only way to avoid this sort of flimsiness is to embrace Absolutism, Conceptualism, and Formalism. I conjecture that it is some – perhaps tacit – recognition of this difficulty that prevents the pluralist alternative from coming into view. But in fact, I will argue, we can avoid the difficulty without embracing these commitments.

The kind of objection to constitutivism I have in mind is exemplified by David Enoch's much discussed essay 'Agency, Shmagency'. He begins by noting that a main attraction of constitutivism, as it is standardly presented, is that it would clarify why certain normative requirements on action are rationally authoritative – why a rational agent cannot coherently ask, concerning such norms, 'Why should I care?' For, constitutivists hold, since the question 'Why should I care?' is asked by an agent considering how to act, and since the relevant norms are implied in the very idea of agency, there can be no genuine question about whether, in the relevant sense of 'should', these norms should be followed. But, Enoch objects, this sort of account of the authority of certain norms simply takes it for granted that one should strive to be what 'an agent' ought to be. If being an agent requires conformity to certain norms to which many people do not actually seek to conform, then it seems coherent to ask 'Why should I be an agent?' But if this question is coherent, then any norms implied by the idea of agency appear to apply to a given subject only conditionally, insofar as she has reason to be an

https://doi.org/10.1017/S1358246117000091 Published online by Cambridge University Press

agent. In the absence of such a reason, it is rationally permissible for her to say 'Agency, Shmagency!' And even if there is such a reason, it presumably cannot itself be grounded in norms implied by the very idea of agency, since it is supposed to ground the applicability of those very norms. So in either case, the constitutivist project of grounding the authority of certain norms in the nature of agency fails.

Now, it is not clear that Enoch's question, 'Why should I be an agent?', is genuinely coherent. Constitutivists generally reply that being an agent is not an attribute that a subject can coherently choose to have or not to have. I *am* an agent, not by choice, but simply by virtue of my nature as a free, self-determining being. Any choice I make will inevitably be the choice of an agent. So, arguably, the question 'Why should I be an agent?' presupposes the possibility of a standpoint that no deliberating subject can actually occupy. If this is right, then any norms implied by the very idea of agency are not just conditionally binding; they are binding on any subject who can confront the question what to do, and the constitutivist project is vindicated.

This reply, however, may seem to be available only to a constitutivist who relies in his argument on the most abstract and minimal conception of agency. For it may seem that being an agent is an inescapable attribute of a subject who confronts the question what to do only if 'being an agent' means nothing more than: being such as to be capable of confronting the question what to do. After all, given any richer conception of agency, it is not a necessary truth that a subject who can act for reasons must be 'an agent' in this sense. So given any richer conception of agency, Enoch's question may seem to regain its grip: why should one strive to be 'an agent' in this richer sense, if it is possible to deliberate, choose, and act without being such a thing? But if Enoch's question gets a grip, then Enoch's dilemma for constitutivism applies. So it appears that a principled constitutivism must aim to ground whatever norms it seeks to vindicate in the very concept of agency, a concept that applies to any conceivable subject capable of acting for reasons. And then a pluralist constitutivism is obviously unacceptable: for if a certain form of agency is supposed to be, not the one-and-only form required in any agent capable of acting for reasons, but one of a plurality of possible forms, then the normative requirements implied by that form will be open to Enoch's challenge, and so the relevant form of agency will not provide the basis for a constitutivist vindication of those norms.

I suspect that a more-or-less explicit concern with heading off such a challenge underlies the widespread assumption that a satisfying

https://doi.org/10.1017/S1358246117000091 Published online by Cambridge University Press

constitutivism must accept Absolutism. Enoch offers an elaborate version of the challenge, but the basic idea is simple. Constitutivists seek to show that certain norms are unconditionally required by observing that anything you choose to do will be an action, and arguing that any action, as such, is subject to certain normative requirements. But, as Millgram puts it,

> The point that anything you do will be an action is likely to be granted only on the thinnest and most minimal reading of what an action is.[19]

Hence it appears that a satisfying constitutivism must assume only the thinnest and most minimal reading of what an action is. But a pluralist constitutivist would be someone who accepts that there are several possible forms of agency, no one of which is necessarily exhibited by all subjects who can act for reasons. Hence a pluralist constitutivist would seem to deprive herself of the cornerstone on which constitutivist arguments are built: the premise that, necessarily, if you act at all, you must do something that is subject to such-and-such normative requirements.

I have described this challenge because I think it provides an intelligible motivation for the assumption that an acceptable constitutivism must be Absolutist. But while this motivation strikes me as intelligible, I think it actually rests on an unsound inference.

It is true that, if there is a plurality of forms of agency, then no one form of agency is necessarily exhibited by all subjects who can act for reasons. Properly understood, however, the constitutivist project need not rely on the claim that there is a form of agency that is necessarily exhibited by any subject who acts for reasons. The constitutivists' claim need not be:

> **The Form:** There is a form of agency such that, necessarily, any subject who acts for reasons will exhibit that form of agency.

It might rather be:

> **My Form:** For any subject who acts for reasons, there is some one single form of agency such that, necessarily, that subject will exhibit that form of agency.

On the latter reading, the starting-point of the constitutivist argument is not the claim that a certain form of agency is inescapable full stop, but that for a given subject, a certain form of agency is inescapable. If this is true, then that agent is inescapable subject to whatever normative requirements are implied by that form of agency, whether or not it is possible for there to be other kinds of

[19] 'Pluralism About Action', 12.

https://doi.org/10.1017/S1358246117000091 Published online by Cambridge University Press

agents subject to other kinds of norms. And that, in effect, is the pluralist's claim: not that there are norms to which any agent, simply *qua* agent, is subject, but that we, *qua* agents of the kind we are, are subject to certain specific norms.

Such a pluralist position will be defensible to the extent that it can be shown that there is a form of agency that is inescapable for us: not an optional refinement that we might choose to exhibit or not to exhibit, but a kind of self-determination that we, being the kind of beings we are, cannot fail to exhibit in our actions. If this condition were satisfied, then we could not coherently ask Enoch's question concerning this form of agency, even if other forms of agency were logically possible.

6. Reason and will: an Aristotelian approach

The preceding section concluded with an abstract description of a possible pluralist constitutivism. But could there actually be a position that met this description? Could we acknowledge that certain fundamental practical norms apply, not to every conceivable agent, but only to agents of a certain form, and yet maintain that, for agents of this form, these norms have an inescapable authority? I think Aristotle's ethics shows the possibility of such a position, and clarifies some of its crucial structural features. My aim in this final section is to sketch – very briefly and programmatically – a reading of Aristotle's ethics along these lines.

It will help to begin by recalling four very familiar features of Aristotelian ethics. First, Aristotle's approach is founded on an appeal to the idea that human beings, as human beings, have a proper work or a function.[20] This, he holds, is what gives determinacy to the inquiry into what human happiness consists in, and thus into how we ought to live. The idea of a human function sounds alien to modern ears, but it becomes clear as the argument proceeds that what Aristotle has in mind is that human beings possess a specific kind of soul, one characterized by certain vital powers or principles, and chiefly, the power of reason. The function of man is to live (and thus to exercise the full complement of these powers) in a manner governed by reason, and Aristotle infers that, since the good of a thing is determined by its function, the human good

[20] Aristotle, *Nicomachean Ethics: Translation, Introduction, Commentary*, ed. Sarah Broadie, trans. Christopher Rowe (Oxford University Press, 2002) I.7.

https://doi.org/10.1017/S1358246117000091 Published online by Cambridge University Press

consists in the excellent performance of those activities that belong to a reason-governed life.

This first point dictates the structure of Aristotelian ethics, but it does not yet determine its content. Although the main divisions of Aristotle's investigation are settled by his conception of the structure of the human soul, what falls within these divisions consists – and this is the second point – of an enumeration of various specific virtues. Aristotle makes no attempt to derive these virtues from his abstract characterization of the human function, and this has seemed to some commentators to make the initial appeal to function superfluous. This reaction, however, depends on a questionable view of the role that the notion of function is supposed to play. Apparently it is not supposed to provide a basis from which to derive particular virtues; Aristotle seems to regard it, rather, as a crucial clarification of what sort of thing a virtue is supposed to be. We might say that it is not the principle of their discovery, but the account of their metaphysical basis – of the kind of fact which can make it the case that a certain putative virtue is a genuine and not a merely pretended one.

Third, Aristotle presents his treatise on ethics as addressed, not to people needing to be convinced that the virtues he enumerates are virtues at all, but rather to those who have been 'properly brought up', brought up in such a way that they already are attracted to these ways of living.[21] Aristotle thus does not seek to address the sort of moral skeptic that Plato confronts in the person of Thrasymachus or Callicles: he merely develops a systematic framework in which to clarify and understand the unity of various goods to which his audience is already unreflectively drawn. And this modesty of justificatory ambition seems closely connected to certain other well-known features of Aristotle's view: his insistence that we should not expect the basic principles of ethics to admit of precise codification;[22] his likening of the understanding of the virtuous person to a capacity for accurate perception rather than for sound reasoning;[23] and his identification of the case-by-case judgment of the practically wise person, rather than some abstract formula, as the standard of right action.[24] All these features of Aristotle's view confirm the observation that he does not expect his general account of the human good to supply the principle of a non-question begging deduction or derivation of the soundness of certain ways of

[21] Ibid., I.5.
[22] Ibid., I.3.
[23] Ibid., II.9.
[24] Ibid., II.6, 1107a1-2, VI.12, 1144a34.

https://doi.org/10.1017/S1358246117000091 Published online by Cambridge University Press

life, or of the right thing to do in particular circumstances. This is not to suggest that he would regard such questions as not open to argument, but he does not seem to aspire to offer a certain sort of foundational argument about such matters.

Finally, the fourth point: although the *Nicomachean Ethics* seeks to give a general account of human virtue, Aristotle is notoriously willing to contemplate the idea that different virtues are appropriate to a man, a woman, a child, and a slave. Whatever else we may make of this, it indicates at least that Aristotle thinks of his official system of virtues as characterizing, not norms applying to all possible agents possessing a rational soul, but norms appropriate to a specific sort of rational agent (namely, the patriarchs of free families living together in a political community). Moreover, in a striking passage, Aristotle remarks that, to the extent that we may think of other species of animals as having practical wisdom, it will be right to say that there is 'a different philosophic wisdom about the good of each species.'[25] The apparent implication is that the practical wisdom he characterizes in the Ethics is, not the only possible form of practical wisdom, but the kind pertaining specifically to human beings. We could express the significance of this by saying that, were there another species of rational animal, the fundamental principles of its practical thought might be quite different from those that govern ours. Perhaps the mode of excellent activity that characterized its reason-governed life would not require courage (think of a rational rabbit); perhaps it might not even be a specifically social rational animal, needing the virtue of justice. At any rate, our practical wisdom is the sound capacity for choice of a certain specific form of agent.[26]

I hope these descriptions of familiar aspects of Aristotle's position resonate with the case made earlier for the possibility of a pluralist constitutivism. Aristotle's ethics is evidently open to the possibility of pluralism, and he certainly does not aim to deduce his system of fundamental norms of action from an analysis of rational agency as

[25] Ibid., VI.7 1141a32-3.

[26] Michael Thompson brought the passage to my attention. Here and throughout I have been influenced by Thompson's work. See his 'Forms of Nature: 'First', 'Second', 'Living', 'Rational' and 'Phronetic'' (unpublished). Also see the following: Matthias Haase, 'Life and Mind', in *The Freedom of Life: Hegelian Perspectives*, ed. Thomas Khurana (Berlin: August Verlag, 2013), 69–109; Jennifer Whiting, 'Hylomorphic Virtue: Cosmology, Embryology, and Moral Development in Aristotle' (unpublished).

https://doi.org/10.1017/S1358246117000091 Published online by Cambridge University Press

such. He does not aim to deduce these norms at all, but simply (i) to give a clarifying description of them to people who already possess an incipient appreciation of their importance, and (ii) to articulate the metaphysical conditions that make it possible for this description to be sound or unsound.

In what sense is this a constitutivist position? I noted earlier that Aristotle's view can be regarded as constitutivist inasmuch as it seeks to ground the fact that we ought to act in certain ways in what we essentially are – where 'what we essentially are' includes, not just that we are rational agents, but that we are specifically human agents. It may seem, however, that this is a constitutivism in name only; for if it does not seek to show that the relevant sort of constitution has inescapable practical authority over those who bear it, and does not seek to deduce specific practical norms from this constitution, in what sense can it claim to ground these norms in our constitution?

At this point, it is important to recall what made the constitutivist approach attractive in the first place. It was, I take it, not its deductive or analytical aspirations *per se*, but rather what the relevant deductions and analyses were supposed to secure: an account of the source of fundamental practical norms that shows how they can be objective but not alien to our will, and that elucidates why their claim to authority over us is not one we can just shrug off, but one to which we are always already committed in the very exercise of our power to choose and act. The basic attraction of constitutivism, then, lies not in its being specially well placed to prove that certain fundamental practical norms apply to us, but in its giving an especially attractive account of what makes it the case that certain fundamental practical norms apply to us. Or to put it another way: the constitutivist's real contribution to metaethics consists, not in offering an account of the epistemology of normativity that would make its claims especially indubitable, but in offering an account of the metaphysics of normativity that would make the nature and authority of these norms particularly transparent.

It seems to me that Aristotle's approach to ethics can lay claim to these advantages of constitutivism in spite of its pluralist, non-analytical, non-deductivist character. What it offers is not a picture of how we can reason to the relevant practical norms; but that is not the basic sense in which the grounding of these norms needs clarification. What needs clarification is what sort of thing a 'fundamental practical norm' might be such that it could have the relevant combination of objectivity, authority, and capacity to move the will. Aristotle sketches a metaphysics that clarifies this by explaining the relation of the relevant

https://doi.org/10.1017/S1358246117000091 Published online by Cambridge University Press

norms to what we essentially are. His account aims to explain why, being the kind of agents we are, and receiving the kind of upbringing that such agents receive in favorable circumstances, we come to find the relevant norms primitively compelling – and also why (i.e., what makes it the case that) we are right to do so. A practically wise Aristotelian agent does not reason from his being (or being committed to being) a human being to specific norms of how he ought to live. His being a human being need not itself come into his practical thought at all, except incidentally. He faces no question about why he should be a human being, or do what human beings, as such, ought to do. It is enough that he is – essentially – a human being, and that this explains why it is sound for the goodness of certain specific ways of living and acting to figure primitively in his thinking about what to do and how to live.

I cannot develop the point here, but I suspect this aspect of Aristotle's ethics reflects much more basic features of his metaphysics. His metaphysics, famously, is hylomorphism: the view that the being of the most basic natural beings (natural substances) consists of a certain type of form being realized in a certain sort of matter. The essence of a natural substance consists, not of its mere form, but of such-and-such-a-form-in-such-and-such matter. In application to human beings, I would argue, this means that our essence is not simply: rational being, or rational animal, but rather: rational-animal-subject-to-such-and-such (specifically human) conditions. And our essence is that sort of being in the absence of we would not be at all. Hence it is not a mode of being one of us might fail to have. The consequence of this point, when brought together with Aristotle's project of explaining good in terms of essence, is that the measure of goodness that flows from our essence is not one that one of us can coherently reject, being what we are. This is not to say that the relevant measure of goodness is unquestionable, but that there is a true and intelligibly well-founded answer to such questions, an answer grounded in what can be called our constitution. We need not accept the substance of Aristotle's ethics to see the structural appeal of such a position. Why should a philosophically significant form of constitutivism claim anything more than this?

University College London
d.lavin@ucl.ac.uk

https://doi.org/10.1017/S1358246117000091 Published online by Cambridge University Press