

OUGHT TO BELIEVE*

Matthew Chrisman
University of Edinburgh

1. INTRODUCTION

We do not have the same sort of control over our doxastic attitudes—belief, disbelief, and suspension of belief—as we do over many ordinary actions. When offered a reward for performing an action, such as raising your hand or turning on the light, you can, at least in normal circumstance, just raise your hand or turn on the light and collect the reward. By contrast, when offered a reward for believing a proposition, such as that the US is still a colony of Britain or that Hell is a bar in Chapel Hill¹, you cannot, at least not in normal circumstances, just believe that the U.S. is still a colony of Britain or that Hell is a bar in Chapel Hill and collect the reward. We can isolate this intuitive contrast in the following *no rewards principle*:

(NRP) No matter how large the reward, S cannot simply decide to believe some proposition p in order to collect that reward.

Because of this intuitive difference, philosophers such as Alston² and Plantinga³ endorse *doxastic involuntarism*:

(DI) The formation of one's beliefs is not within one's direct voluntary control.⁴

* For helpful feedback on previous versions of this paper, I would like to thank John Broome, Campbell Brown, Andy Clark, Graham Hubbs, Hilary Kornblith, Ram Neta, Duncan Pritchard, Michael Ridge, Jay Rosenberg, Russ Shafer-Landau, Piers Turner, and audiences at the University of Aberdeen and the University of Edinburgh.

¹ The first example is from W. Alston, "The Deontological Conception of Epistemic Justification," *Philosophical Perspectives*, 2 (1988): 257-299. I also include the second example because it illustrates that the issue is not whether you can voluntarily believe a false proposition in order to collect a reward but whether you can voluntarily believe a proposition for which you have countervailing evidence or even no evidence.

² See *Ibid.* and *Beyond Justification: Dimensions of Epistemic Evaluation* (Ithaca: Cornell University Press, 2005).

³ See *Warrant: The Current Debate* (New York: Oxford University Press, 1993).

⁴ In "Deciding to Believe," in *Problems Of The Self: Philosophical Papers, 1956-1972* (Cambridge: Cambridge University Press, 1973), p. 136-147, B. Williams endorsed the stronger principle that the formation of one's beliefs cannot be within one's voluntary control. He thought that this is an implication of the idea that "belief aim at truth". There has been much discussion of what to make of this idea and the conclusion Williams draws from it. See especially J. Bennett, "Why is Belief Involuntary?" *Analysis* 50:2 (1990): 87-107, D. Scott-Kakures "On Belief and the Captivity of the Will," *Philosophy and Phenomenological Research*, 54:4 (1993): 77-103, D. Radcliffe "Scott-Kakures on Believing at Will," *Philosophy and Phenomenological Research*, 57:1 (1997): 145-151, D. Velleman "On the Aim of Belief,"

In this way, belief formation seems to be much more like digestion than ordinary actions such as raising one's hand or turning on the lights.⁵ As with digestion, one may have *indirect* control over the formation of beliefs in that there are actions within one's direct voluntary control that have foreseeable effects on what one believes. But, in both cases, we seem to lack the sort of control that we have over ordinary actions.

However, it also seems clear that many statements about what we ought to believe, disbelieve or suspend belief about are true. For example, you ought to believe that you are reading this paper right now. We can isolate this intuition in the following *true doxastic oughts principle*:

(DOP) At least some sentences of the form 'S ought to believe p' are true.

Because of this intuitive idea, philosophers such as M. Steup⁶ and S. Ryan⁷ endorse what we might call *epistemic deontology*⁸:

(ED) Beliefs are proper subjects of epistemic oughts.

To many who are impressed by the idea that 'ought' implies 'can', epistemic deontology has seemed to be in tension with doxastic involuntarism. This has led to a standoff between doxastic involuntarists and epistemic deontologists. Doxastic involuntarists argue that epistemic

The Possibility of Practical Reason (Oxford; New York: Clarendon Press Oxford, 2000), N. Shah "How Truth Governs Belief," *The Philosophical Review*, 112:4 (2003): 447-82, and N. Shah and D. Velleman "Doxastic Deliberation," *The Philosophical Review*, 114, 4 (2005): 497-534. For my purposes here, however, the weaker form of doxastic involuntarism stated in the text is strong enough, and here I will remain neutral about its relation to the idea that belief aims at the truth.

⁵ Raising one's hand is an action over which we have what Alston refers to as "basic voluntary control" because it is something we can just do. Turning on the light is an action over which we have what Alston refers to as "nonbasic immediate voluntary control" because it is something we can do right away by doing something else (such as, flipping the switch). By in large, these details will not matter here, but it is worth mentioning that Alston claims that we have neither sort of voluntary control over our beliefs, although he allows that we have what he calls "nonbasic indirect voluntary control" of a very weak sort over our beliefs, comparable to the sort of control we have over our blood pressure. R. Feldman "The Ethics of Belief," *Philosophy and Phenomenological Research*, 60:3 (2000): 667-695, argues that there are some beliefs over which we exercise nonbasic immediate voluntary control. His example is the belief that the lights are on in his office. If offered a high reward to believe this, he could collect the reward by turning on the lights in his office and thereby coming to have the belief right away. Alston might respond that this is indirect rather than direct nonbasic voluntary control, but I think the line between the two is vague. What is important here is that (as Feldman recognizes) the sort of voluntary control over beliefs that his example illustrates is something we have over only a small class of potential beliefs. Basically it is just the beliefs for which we have the power to make their propositional contents true. For the purposes of my discussion here, this class of beliefs may be bracketed, since the sorts of ordinary doxastic oughts that generate the tension between doxastic involuntarism and epistemic deontology far outstrips this class of beliefs.

⁶ "Doxastic Voluntarism and Epistemic Deontology" *Acta Analytica*, 15, 24 (2000): 25-56, and "Doxastic Freedom," *Synthese*, forthcoming.

⁷ "Doxastic Compatibilism and the Ethics of Belief," *Philosophical Studies*, 114 (2003): 47-79.

⁸ This is a perhaps somewhat restricted way to use the term 'epistemic deontology', which a comparison to ethical deontology reveals: ethical deontologists, consequentialists, and virtue theorists all agree that actions are proper subjects of ethical oughts; what they disagree about is the source or justification for these ethical oughts. So, epistemic deontology might rightly be thought to involve something more to distinguish it from epistemic consequentialism, and epistemic virtue theory. For more on a pure form of epistemic deontology, see P. N. Turner "Epistemic Deontology and the Consequentialist Consensus," Masters Thesis presented to the faculty of the University of North Carolina, at Chapel Hill, 2004. However, whatever more is required, the view will include a commitment to the idea that beliefs are proper subjects of doxastic oughts, and the cogency of this idea is what will be under discussion here.

deontology must be false, and are thus forced to explain away the intuitions behind the true doxastic oughts principle, while epistemic deontologists argue that doxastic involuntarism must be false, and are thus forced to explain away the intuitions behind the no rewards principle.

My primary purpose in this paper is to sketch a theory of doxastic oughts that achieves a satisfying middle ground between the extremes of rejecting epistemic deontology because of the no rewards principle and rejecting doxastic involuntarism because of the true doxastic oughts principle. The key will be appreciating the fact that not all true oughts require direct voluntary control. I will construct my account as an attempt to surpass other accounts (especially those due to Feldman and Kornblith) in this vein. The new idea (in a telegraphic slogan) is that doxastic oughts are what Sellars called “rules of criticism”, which are logically distinct from but also interestingly connected to “rules of action”. The distinction provides a way to understand the phrase “ought to believe” which is consistent with both doxastic involuntarism and epistemic deontology; the connection provides a novel way to incorporate a believer’s epistemic community into our understanding of the scope of epistemic obligations.

2. REJECTING DOXASTIC INVOLUNTARISM: A MISTAKEN REACTION TO THE TENSION

To begin I want to consider two arguments against doxastic involuntarism because I think the way in which they fail is instructive. The first can be extracted in Ryan’s response to premise 2 of what she calls “The Anti-Voluntarist Argument”, which goes:

1. If we have any epistemic obligations, then doxastic attitudes must sometimes be under our voluntary control.
2. Doxastic attitudes are never under our voluntarily control.
3. We do not have any epistemic obligations (1), (2).⁹

As we have already seen, doxastic involuntarism is motivated in large part by the no rewards principle, which seems to illustrate that we do not have the power simply to decide whether to believe or disbelieve *p*. Ryan directly denies this: “if I make a decision to switch doxastic attitudes, I can adjust my doxastic attitudes accordingly.”¹⁰ But what then does she say about the intuitive examples which motivate the no rewards principle? Ryan thinks they do not undermine doxastic voluntarism because they are not cases where one is given good reason to change one’s attitudes. She writes, “My doxastic decisions are guided by what seems to me to be good evidence (or whatever it is that motivates me).”¹¹ She grants that “Without new evidence, or a new appreciation of my evidence, I cannot typically just decide to switch propositional attitudes”¹², but she insists that, “...once I have a reason to decide to change attitudes, I can and I do. This limitation is not a limitation of my doxastic freedom and it is not a good reason to accept premise (2) of [the Anti-Voluntarist Argument].”¹³

However, there seems to me to be an important equivocation in this argument. Ryan talks of the need for “evidence”, “reason to decide to change attitudes”, and, cryptically, “whatever it is that motivates me” in order to decide to change propositional attitudes. Yet, surely the offer of a large reward to believe that the U.S. is still a colony of Great Britain is a pretty good “reason to decide to change attitudes”? It seems that, if such decision were in Ryan’s

⁹ *Op. cit.*, p. 48. Ryan actually thinks both premises are false. Below I will consider some reasons one might reject premise 1, but that is, of course, consistent with doxastic involuntarism. Compare also Steup, *op. cit.*, 2000, p. 26, and Feldman, *op. cit.*, 2000, p. 669, where similar arguments are criticized.

¹⁰ *Op. cit.*, p. 65.

¹¹ *Ibid.*

¹² *Ibid.*, p. 66.

¹³ *Ibid.*

power, the reward surely would “motivate” her. The fact that it does not is what is supposed to show that such decision is not within her power.

Hieronymi¹⁴ makes a distinction that is helpful for diagnosing the error in Ryan’s argument. She distinguishes between what she calls “constitutive reasons” and “extrinsic reasons” for a belief. Constitutive reasons for the belief that *p* are reasons that bear on the question whether *p*; to find them convincing just is to believe that *p*. By contrast, extrinsic reasons for the belief that *p* are reasons that count in favor of believing *p* independently of whether *p*. They are reasons that bear rather on the question whether it would be good to believe that *p*; finding them convincing does not constitute believing that *p*. In light of this distinction, Ryan’s claim that “My doxastic decisions are guided by what seems to me to be good evidence (or whatever it is that motivates me)” can be seen to run together two different kinds of reasons for belief. Evidence provides constitutive reasons for belief, while other considerations provide (if anything) extrinsic reasons for belief.

When the involuntarist claims that believing that *p* is not the sort of thing that one can voluntarily decide to do, I think this should be understood as the claim that beliefs are not responsive to practical reasons in the same way that actions are. That is, beliefs are not responsive to what Bennett¹⁵ described as reasons that bear on what to make true as opposed to reasons that bear on what is true. Practical reasons are typically a set of considerations that show an action to be somehow choice worthy. For example, getting a large reward could be a reason to do what it takes to get the reward. However, such considerations do not typically bear on the question whether a particular proposition is true, so they are not—in Hieronymi’s terminology—constitutive reasons for belief. For instance, the fact that you’ll receive a large reward for believing that the US is still a colony of Britain clearly does not bear on the question whether the US is still a colony of Britain, so the reward is not a constitutive reason for that belief. What the reward does bear on is whether it would be worthwhile to have that belief independently of whether it is true. This means that such considerations are—in Hieronymi’s terminology—extrinsic reasons for belief. So, if choice requires practical reasons and practical reasons are extrinsic reasons for belief, it would seem that the ability to voluntarily decide to believe some proposition *p* requires that one be capable of believing *p* on the basis of extrinsic reasons for belief. But, what the no rewards principle shows is precisely that we cannot typically form a belief for extrinsic reasons (even if we can, for such reasons, perform an action that has as a consequence that one comes to have certain beliefs).¹⁶ When offered a large reward for believing that the US is still a colony of Britain, one cannot do it, though one may be capable of doing something else, like undergoing shock therapy (or whatever), designed to induce the belief.

So I think Ryan’s argument fails in its conflation of constitutive reasons and extrinsic reasons for belief. The former do not count as reasons to decide to do something because they are not practical reasons, and, while the latter may be practical reasons for deciding on some action, they do not motivate belief. However, it is important to notice that premise 2 of the anti-voluntarist argument does not say that we never have the power to voluntarily decide on our doxastic attitudes; rather it says that our doxastic attitudes are never under our voluntary control. This opens up space for an alternative strategy for undermining that premise, which in

¹⁴ “The Wrong Kind of Reason,” *The Journal of Philosophy*, 102 (2005): 437-57. For a different use of the same distinction, see also P. Hieronymi “Controlling Attitudes,” *Pacific Philosophical Quarterly*, 87 (2006), p. 51-52.

¹⁵ *Op. cit.*

¹⁶ See Hieronymi, *op. cit.*, 2006, p. 52-60, for a plausible explanation of why extrinsic reasons cannot directly motivate belief. The details are not important here, but what is important is that the no rewards principle shows that extrinsic reasons cannot motivate belief, which seems to undermine doxastic voluntarism.

turn generates a different sort of argument against involuntarism. To see this, notice that the voluntarist could concede to the involuntarist that we do not have the power to voluntarily decide on our doxastic attitudes but then argue that this power is not a necessary condition on voluntary control. If it is not necessary for voluntary control, then perhaps we have voluntary control over our beliefs even though we lack the ability to voluntarily decide on our doxastic attitudes. Given the way it is formulated above, this possibility would clearly undermine doxastic involuntarism.

Assuming that voluntary decision is a kind of choice, this argument against involuntarism is akin to Steup's response to what he calls "The Argument from Choice", which goes:

- P1 One's ϕ ing is free iff one could have chosen not to ϕ .
- P2 I could have chosen not to have a cup of coffee.
- P3 I could not have chosen not to believe that cats are mammals.
- C Whereas my having a cup of coffee is free my believing that cats are mammals is not.¹⁷

Locating the misstep in Ryan's argument helped to highlight the importance of the distinction between constitutive and extrinsic reasons for belief, but it left open the possibility that doxastic involuntarism is false because voluntary control doesn't require ability to decide voluntarily. Likewise, locating the misstep in Steup's response to this new argument will help us to see why this possibility doesn't undermine doxastic involuntarism.

Steup considers three conceptions of what it means to say that one could have chosen not to ϕ :¹⁸

- C1 I ϕ d but I could have chosen not to ϕ iff I ϕ d but had I a reason not to ϕ I would not have ϕ d.

He argues that on this conception, P3 is false. He writes, "Had I had a reason not to believe that cats are mammals, I would not have believed it...doxastic attitudes are no less reason responsive than our actions."¹⁹ So, if a proponent of DI rests his case on the idea that voluntary control in doing ϕ requires the power to choose not to ϕ , he will need a conception of choice different from C1. Next he considers:

- C2 I ϕ d but could have chosen not to ϕ iff I ϕ d but had I intended not to ϕ I would not have ϕ d.

Regarding this conception, he concedes that "our doxastic attitudes do not respond to intentions, so doxastic attitudes never meet the condition C2 states"²⁰; however, he argues that responsiveness to intentions is not a necessary condition on something's being done freely. He writes, "Assessing the freedom of actions calls for one yardstick, assessing the freedom of doxastic attitudes for another. It's a mistake to think that the freedom of actions and the freedom of doxastic attitudes can be gauged using one single yardstick."²¹ Finally there is:

- C3 I ϕ d but could have chosen not to ϕ iff I ϕ d but had I decided not to ϕ I would not have ϕ d.

And regarding this, he argues, in effect, that it collapses into C1. If deciding is just deliberating by weighing reasons in order to reach a conclusion, then he thinks we surely can weigh reasons

¹⁷ *Op. cit.*, forthcoming, p. 20 of ms.

¹⁸ These come from *ibid.*, p. 20-21 of ms.

¹⁹ *Ibid.* Of course, they seem to be responsive to a different kind of reason, but more on this below.

²⁰ *Ibid.*

²¹ *Ibid.*

for belief in order to reach a decision. This leads to the view that “Weighing my practical reasons, I might decide not to perform a certain action. Likewise, weighing my epistemic reasons, I might decide not to adopt a certain doxastic attitude. If ‘deciding not to φ ’ is thus understood broadly, P₃ is false.”²² Of course, one might insist that deciding is not mere responsiveness to reasons but instead the formation of an intention, and, since beliefs do not appear to be responsive to intentions, they fail to count as voluntary according to C₃. However, I think Steup would argue that this response, in effect, collapses C₃ into C₂, to which his response, again, is that responsiveness to intentions is not a necessary condition for voluntary control.

That is Steup’s response to the Argument from Choice. On any of the three conceptions of choice, he thinks the argument fails to be sound. With respect to the debate between epistemic deontology and doxastic involuntarism, the thrust of Steup’s argument here is that thinking of beliefs as proper subjects of broadly epistemic oughts presupposes only that some beliefs are free and not that beliefs are things that we can choose in the narrow sense of choice embodied in C₂. So, if, in the idea that ‘ought’ implies ‘can’, ‘can’ means ‘can choose’ in the sense of C₂, Steup will say that this idea is simply mistaken. To think otherwise, he contends, is to measure the voluntariness of doxastic attitudes by the wrong “yardstick.”²³

However, this raises the question: by what yardstick *are* we to measure the voluntariness of doxastic attitudes? And Steup starts us in an apparently promising direction: according to him, it has something to do with responsiveness to *epistemic* reasons. A walk in the park is voluntary if I do it for practical reasons which do not involve coercion, manipulation, paranoia, etc. Likewise believing that cats are mammals is voluntary if I believe that for epistemic reasons which do not involve wishful thinking, prejudice, hypnosis, etc. This both distinguishes doxastic attitudes from actions and locates them under one genus of potential freedom—viz., responsiveness to the appropriate sort of reasons. Moreover, it helps to explain why beliefs formed by wishful thinking, prejudice, hypnosis, etc. are not within our voluntary control: they are not responsive to epistemic reasons.

I think we should grant that Steup’s notion of responsiveness to reasons provides a way to distinguish between two significant classes of doxastic attitudes—“free” and “unfree”. What I would not grant, however, is that this distinction can be used to resolve the tension between doxastic involuntarism and epistemic deontology. For, although they are free from irrational influence, “free” doxastic attitudes are not things over which we exercise direct voluntary control. It is because of this that the no rewards principle is so intuitively attractive. Even though our doxastic attitudes are sometimes responsive only to epistemic reasons and thus “free”, believing or disbelieving a particular proposition is not the sort of thing that it is coherent to offer to reward someone to do.

Why not? As far as I can tell, Steup’s best response to this question is to explain the intuitive difference between doxastic attitudes and ordinary actions, on which the no rewards principle turns, as a reflection of the difference between the types of reasons to which doxastic attitudes and ordinary actions are usually responsive. Both can be “free” because both are responsive to reasons, but the former are, when free, responsive to epistemic reasons while the latter are, when free, responsive to practical reasons. Yet, as we saw above, the offer of a reward might provide a practical reason but it surely provides no epistemic reason; this is because such rewards are merely extrinsic reasons for the belief, while epistemic reasons are constitutive reasons. And so I suspect opponents of epistemic deontology will take this response as grist to their mill. For their problem with epistemic deontology should not be its dependence on the idea

²² *Ibid.*

²³ *Ibid.*

that beliefs are free on this sense of “free” but rather the fact that it implies that ‘ought’ can apply to things over which we lack voluntary control.

Steup does anticipate something like this worry when he argues against what he calls “Practical Reasons Chauvinism” which is the view that, “Whereas responsiveness to practical reasons grounds freedom, responsiveness to epistemic reasons does not.”²⁴ However, the point I am making here is that even if we grant that responsiveness solely to epistemic reasons makes doxastic attitudes free, this is different from the sort of control opponents of epistemic deontology think is necessary for true doxastic oughts. About apparently free actions and apparently free beliefs, Steup writes,

...there is a reason-responsiveness in either case. Thus there is a strong *prima facie* case for thinking there is voluntary control in either case. If opponents of doxastic freedom wish to resist this line of reasoning, they need to justify the chauvinistic premise that responsiveness to epistemic reasons does not count as a freedom-grounding kind of responsiveness.²⁵

But I think Steup conflates freedom and voluntary control in this passage. The involuntarist can grant that doxastic attitudes can often be free in that they are often *free from* irrational influence and responsive to epistemic reasons. There need be no chauvinism about types of freedom here. But the involuntarist will insist that, since doxastic attitudes are not responsive to practical reasons, they are not under our direct voluntary control. And—the thought goes—since ‘ought’ implies ‘can’, demonstrating lack of voluntary control is all that is necessary to undermine epistemic deontology.

3. REJECTING “OUGHT IMPLIES CAN”: A STEP TOWARDS A SOLUTION

So far, we have been considering defenses of epistemic deontology that rest on rejecting doxastic involuntarism. But, since I think these defenses are not successful and because I would like to develop an account of doxastic oughts that is consistent with both doxastic involuntarism and epistemic deontology, I now want to reconsider the popular idea that ‘ought’ implies ‘can’, which seems crucial to the involuntarists’ argument against epistemic deontology. In the literature about this, there are a variety of ostensible counter examples and attempts to hone the idea to avoid them.²⁶ For example, it seems true that you ought to pick up your friend on time, even if, because you left too late, you can no longer do so. Also, it seems true that your friend ought not to steal, even if, because she’s a kleptomaniac, she cannot help herself. However, in my view, these sorts of counterexamples and the proposed fixes are a red-herring in the debate between proponents of doxastic involuntarism and epistemic deontology. This is because they are examples of *actions* that all participants to that debate would agree *can* in principle be under our direct voluntary control even if there are cases when they are not. The vexing thing about doxastic attitudes, however, is not that there are some situations where we ought to have a particular doxastic attitude even though we do not happen to exercise voluntary control over them in that situation; rather the vexing thing is that the no rewards principle seems to show that doxastic attitudes are never under our direct voluntary control but it still seems as if we ought to have some of them. How can this be? Merely undermining the general ‘ought’ implies ‘can’ principle with a counterexample based on actions cannot answer this question.

²⁴ *Ibid.*

²⁵ *Ibid.*, pp.18-19 of ms.

²⁶ For discussion of several epicycles, see Ryan, *op. cit.*, drawing on M. Stocker, “‘Ought’ and ‘Can’” *Australasian Journal of Philosophy*, 49, 3 (1971): 313-317; and W. Sinnott-Armstrong, “‘Ought’ Conversationally Implies ‘Can’,” *The Philosophical Review*, 93, 2 (1984): 249-261.

This point is relevant for evaluating two prominent proposals for understanding doxastic oughts in the face of the tension between the no rewards principle and the true doxastic oughts principle. The first comes from Feldman, who argues that doxastic oughts are akin to *role oughts*.²⁷ As examples he suggests that, “Teachers ought to explain things clearly. Parents ought to take care of their kids. Cyclists ought to move in various ways.”²⁸ In the present discussion, the significance of these is that role ‘ought’s seem not to imply ‘can’s.²⁹ An incompetent teacher ought to explain things clearly but she cannot; an incapable parent ought to take care of his kids, but he cannot; an untrained cyclist ought to move in various ways, but she cannot.³⁰ If doxastic oughts were similar, then the fact that we do not have direct voluntary control over our doxastic attitudes would not undermine epistemic deontology. This is why Feldman suggests that we often are in the role of a believer. He writes, “...we form beliefs in response to our experiences in the world. Anyone engaged in this activity ought to do it right. In my view, what they ought to do is to follow their evidence (rather than wishes or fears). I suggest that epistemic oughts are of this sort.”³¹

However, a serious problem with this proposal is that, unlike Feldman’s examples of role oughts, doxastic oughts seem to be categorical. Even if you do not want to believe the truth about what you are doing right now, you ought to believe that you are reading this paper right now.³² As Kornblith puts the point, in “the epistemic case,” by contrast to optional roles like teacher, parent, or cyclist, “we not only want to say that if someone wants to be a good believer, he or she should believe in certain ways; we also wish to endorse the claim that individuals ought, without qualification, to believe in those ways which, as a matter of fact, flow from good performance of the role of being a believer.”³³ It seems that Feldman’s only means for explaining this contrast is the fact that, unlike the roles of teacher, parent, and cyclist, we have no choice about whether to take on the role of being a believer. However, this is a bad explanation. The kleptomaniac may also have no choice about whether to take on the role of a thief, but we would not want to say that she categorically ought to steal. Exactly the opposite: she categorically ought

²⁷ *Op. cit.*, p. 675.

²⁸ *Ibid.*

²⁹ Later in Feldman “Modest Deontologism in Epistemology,” *Synthese*, forthcoming, he makes a compatible proposal, suggesting that “S ought to believe that p” should be analyzed as “Believing that p is the epistemically appropriate response to S’s evidence.” He takes this to be a way of spelling out the idea that in one’s role as a believer one ought to follow one’s evidence. Whatever way we spell out the details, I still think Feldman’s proposal falls prey to the criticisms I go on to articulate above.

³⁰ Campbell Brown has suggested to me in conversation that Feldman’s thought here involves a scope confusion. It is true that not all teachers can explain things clearly, but this undermines only a narrow scope reading of “Teachers ought to explain things clearly”. The narrow scope reading is “For all x, if x is a teacher then ought (x explains things clearly).” The wide scope reading is “For all X, ought(if X is a teacher, then X explains things clearly)”, and Brown’s idea is that this is not undermined by the fact that not all teachers can explain things clearly. The ones who cannot could satisfy this obligation by quitting their jobs. There are naturally some questions about whether this wide-scoping move is fully defensible. But, in any case, it seems wrong to me to think of doxastic oughts as role oughts for reasons that emerge more clearly in the text below.

³¹ *Ibid.*

³² For a similar point, compare T. Kelly, “Epistemic Rationality as Instrumental Rationality: A Critique,” *Philosophy and Phenomenological Research*, 64 (2003).

³³ “Epistemic Obligation and the Possibility of Internalism” in A. Fairweather and L. Zagzebski, eds. *Virtue Epistemology: Essays on Epistemic Virtue and Responsibility*, (New York; Oxford: Oxford University Press, 2001), p. 237. Kornblith might not want to call these “categorical” oughts but rather a sort of hypothetical ought where the hypotheticality is cancelled for any agent with any desires—for the full view, see his *Knowledge and its Place in Nature*, (Oxford; New York: Clarendon Press; Oxford University Press, 2002), ch. 5.

not to steal. So the mere fact that it is our “plight” to play certain roles cannot explain the apparent categoricity of doxastic oughts.

Kornblith has an alternative proposal, which I think provides a crucial insight. He suggests that what Feldman is right about is that some oughts come from evaluations of what counts as good performance; and the cogency of such oughts clearly does not require voluntary control to do what they prescribe or proscribe, as illustrated by role oughts deriving from roles such as teacher, parent, and cyclist. But, according to Kornblith, this is also true of *ideals*. He writes,

An appropriate human ideal must in some ways be responsive to human capacities. Ideals are meant to play some role in guiding action, and an ideal that took no account of human limitations would thereby lose its capacity to play a constructive action-guiding role. At the same time, our ideals cannot be so closely tied to what particular individuals are capable of that we fail to recognize that some individuals at some times are incapable of performing in ideal ways. There is a large middle ground here, and it is here that reasonable ideals are to be found.³⁴

So, for example, respecting the legitimate property of others is plausibly thought to be a moral ideal, and from this ideal, it probably follows that one ought not to steal. The fact that, for example, your kleptomaniac friend cannot help but to steal does not undermine the fact that she ought not to steal.

Kornblith thinks this helps to preserve epistemic deontology in the face of the no rewards principle. He writes, “once we recognize that our ideals must lie somewhere within this large middle ground, we see that the defensibility of the oughts that flow from our epistemic ideals does not require the level of voluntary control over our beliefs that Alston and Plantinga insist upon”(Ibid.). But because these oughts flow from ideals rather than roles, their normative force can be categorical unlike the role oughts cited by Feldman. Kornblith writes, “Although the role of being a slave might be performed well, it is perfectly reasonable to suppose that being a slave, whether the role be performed well or badly, is no part of any acceptable human ideal. It is for this reason that oughts which flow from human ideals have a degree of normative force that is not shared by role oughts.”³⁵

I think Kornblith is right that doxastic oughts derive from our epistemic ideals. And it is certainly correct that human ideals may transcend particular humans’ capacities without undermining the oughts that derive from them. The moral ideal of respecting others’ property underwrites the moral proscription, “One ought not to steal,” even if your kleptomaniac friend cannot help herself. Nevertheless, I do not think Kornblith’s comparison to moral ideals is very helpful for addressing the apparent tension between the true doxastic oughts principle and the no rewards principle. This is because the problem the no rewards principle was supposed to cause for epistemic deontology is not that true doxastic oughts transcend *particular* humans’ capacities to form beliefs but rather that forming beliefs does not seem to be the sort of thing that *anyone* has direct voluntary control over. Again, the upshot here is that action-types that provide ostensible counterexamples to the idea that ‘ought’ implies ‘can’ are not suitable as models for doxastic oughts in the fact of the no rewards principle, which seems to show that beliefs are never actions over which we exercise direct voluntary control.

4. A BETTER ACCOUNT

³⁴ *Op. cit.*, 2001, p. 238.

³⁵ *Ibid.*, p. 238-239.

In different ways, my criticism of Ryan and Steup, on the one hand, and Feldman and Kornblith, on the other, rest on the same underlying thought: There is an important difference between actions like raising one's hand and turning on the lights and states of mind like belief and disbelief. To show that categories like "freedom" and "reason-responsive" can apply to both, is not yet to undermine doxastic involuntarism nor is it to explain (away) the no rewards principle. And to show that there are cases where it looks like someone ought to perform some action that she cannot is not yet to show doxastic involuntarism to be compatible with epistemic deontology nor is it to explain the no rewards principle. For the no rewards principle is an illustration not of the fact that *sometimes some people* cannot do what they ought to do but rather the fact that believing is not, in the relevant sense, an action that *anyone* can *ever* voluntarily perform.

So, in developing a more satisfactory view of doxastic oughts, I think it is helpful to notice that there are all sorts of oughts that apply not to actions but to states. For instance, your doctor might advise that "You ought to have lower blood pressure," or the clock repairman might point out that "The clock chimes ought to strike on the quarter hour," and the camp counselor might tell her campers that "All the beds ought to be made every morning." These oughts are oughts that, when true, seem to be true despite the fact that your blood pressure, the clock's chiming, and the beds being made are not, properly speaking, things that are in someone's direct voluntary control. This is not because we (or some of us) happen to lack control over a particular sort of action, rather it is because these things are not actions at all. They are ways of *being* rather than actions that some agent *does*.

This encourages the following distinction: we seem to have 'ought-to-do's, on the one hand, and 'ought-to-be's, on the other—or, what Sellars termed "rules of action" and "rules of criticism".³⁶ Only the former seem to presuppose any sort of control on the part of the subject of the ought.

Now, this distinction may seem to be unavailable if we follow the standard view in deontic logic and treat 'ought' as a unary modal operator applying to propositions.³⁷ On a very schematic version of this view, all ought-statements with a particular subject can be transformed into impersonal constructions beginning with "It ought to be the case that..." in a way that elides the distinction between 'ought-to-do's and 'ought-to-be's. For example, "Eric ought to pick up Ellie on time" and "The clock ought to strike on the quarter hour" would be transformed into "It ought to be the case that Eric picks up Ellie on time" and "It ought to be the case that the clock strikes on the quarter hour." Treating 'ought' as this kind of unary modal operator is helpfully simplifying for developing the semantics and proof theory for deontic logic; however, the problem with this in the present context is that this way of proceeding seems to wash away the distinction between oughts that might reasonably be thought to presuppose voluntary control of their subjects and those which do not.³⁸ This is why I want to resist the idea that ought is just a

³⁶ "Language as Thought and as Communication," *Philosophy and Phenomenological Research*, 29 (1969): 506-527. For a similar distinction, compare L. Humberstone, "Two Sorts of 'Ought's,'" *Analysis*, 32 (1971): 8-11.

³⁷ See McNamara, Paul, "Deontic Logic", *The Stanford Encyclopedia of Philosophy* (Spring 2006 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/spr2006/entries/logic-deontic/>>.

³⁸ There are some attempts to modify the standard view to include a place for the agent in the proposition to which the ought-operator applies. For instance, by deploying a 'stit' ("see to it") device in the statement of the proposition to which an ought applies, we can create a spot for an agent. For example, "It ought to be the case that Eric picks up Ellie" could be construed as "It ought to be the case that Eric sees to it that he picks up Ellie on time," in order to preserve reference to agency. See J. Horty, *Agency and Deontic Logic* (Oxford: Oxford University Press, 2001). See also J. Broome, "Normative Requirements," *Ratio*, 12, 4 (2000): 398-419, and R. Wedgwood, "The Meaning of 'Ought'," *Oxford Studies in Metaethics*, 1 (2006): 127-160, for other attempts to modify the standard view to preserve reference to agency. I will not discuss

unary modal operator applying to propositions.³⁹ However, that does not commit me to saying that ‘ought’ is never a unary modal operator on propositions. ‘Ought’ clearly has a nonnormative use in sentences such as “Given how the sky looks tonight, it ought to rain tomorrow.” Moreover, some uses of ‘ought’ may be merely “political” as Sidwick⁴⁰ calls them (to distinguish them from “agential”).⁴¹ The idea is that some things ought to be the case because they are generally desirable, although no particular agent ought to do anything to make things this way. For example, perhaps the party ought to start early (so as not to disturb the neighbors later), but no one in particular is responsible for starting the party early. The nonnormative ought and the political ought may be plausibly interpreted as unary modal operators on propositions, but, even so, I think it would be wrong to treat ‘ought’ always in this way. The distinction between rules of criticism and rules of action is supposed to help distinguish between two kinds of oughts that are both not unary modal operators.

Insisting on this distinction is not meant to rule out the possibility that rules of criticism and rules of action are logically related. In fact, Sellars boldly claimed that, “...though ought-to-be’s are carefully to be distinguished from ought-to-do’s they have an essential connection with them. The connection is, roughly, that ought-to-be’s imply ought-to-do’s.”⁴² It will prove helpful briefly to explore what Sellars has in mind. The general form of the implication is material (rather than formal)⁴³ and involves a heavy *ceteris paribus* clause, but I take the core idea to be that statements of the form,

X’s ought to be in state φ ,

where these are not merely nonnormative or political uses of ‘ought, materially imply statements of the form,

(Other things being equal and where possible) one ought to bring it about that X’s are in state φ .

Sellars puts this distinction to crucial use in a general account of linguistic activity as through and through rule-governed, which is nonetheless compatible with denying that all (or even a significant portion of) linguistic activity should be thought of as actions performed in accordance with particular rules. The leading idea is to construe most linguistic rules as rules of criticism rather than rules of action. He writes, “To be a language user is to conceive of oneself as an agent subject to rules.”⁴⁴ But the value of the distinction between rules of action and rules

these views in any detail here. As will emerge below, although I think we need to preserve reference to a not merely grammatical subject of some oughts, this subject won’t always be an *agent*.

³⁹ For other, more grammatical arguments for resisting the standard view in deontic logic, see M. Schroeder, “Do Oughts Take Propositions” (manuscript).

⁴⁰ *The Methods of Ethics*, 7th edition (London: Macmillan, 1907), p. 34.

⁴¹ See also B. Williams, “Ought and Moral Obligation,” in his *Moral Luck* (Cambridge: Cambridge University Press, 1981); and R. Wedgwood, *op. cit.*, for discussion of this distinction. Wedgwood suggests a modification of the standard view in order to preserve the agential nature of some oughts while still construing them as unary modal operators.

⁴² *Op. cit.*, p. 508.

⁴³ That is to say that it is not merely in virtue of the logical form of the statements that an inference from one rule to the other is correct, rather it is in part because of the nature of the concepts deployed in the statements that such an inference is correct. For example, the inference from “x is a triangle” to “x has three sides” is materially correct. (Compare Sellars “Inference and Meaning,” *Mind*, 62, 247 (1953): 313-338, on the materially correct inference from “Pittsburgh is to the west of Princeton” to “Princeton is to the east of Pittsburgh.”)

⁴⁴ *Op. cit.*, 1969, p. 513.

of criticism, according to him, is that “one can grant this without holding that all meaningful linguistic episodes are *actions* in the conduct sense, and all linguistic rules, rules for doing.”⁴⁵

Here is not the place to evaluate the complex role of rules in Sellars’ general theory of linguistic activity, but I do want to take away one lesson from it. Although the paradigmatic rules of criticism apply to things that are not agents—such as clocks and beds—there is nothing to prevent rules of criticism from applying to agents as they are in various states. Moreover, this does not impinge on the way that rules of criticism imply interpersonal rules of action. So, for example, Sellars suggests that a possible rule governing basic perceptual reports is that,

One ought to respond to red objects in sunlight by uttering or being disposed to utter ‘this is red’ under appropriate prompting.⁴⁶

Although this alleged rule of criticism applies to speakers of a language and to those who are in the process of learning a language, it is not to be confused with the (merely putative) rule of action,

One ought to say ‘this is red’ in response to red objects in sunlight under appropriate promoting.

In his view, this rule of action could be understood and so voluntarily followed only by those who already have the concepts of red, object, sunlight, and what it is to say something. So, importantly, Sellars thinks this could not constitute a rule language learners genuinely follow in acquiring the concept of red (or, for that matter, the concepts object, sunlight, and what it is to say something).⁴⁷ But, in Sellars’ view that does not mean that the rules of criticism that govern linguistic activity have no connection to rules of action, for, in his view, rules of criticism imply rules of action. In particular, he suggests that the above ought-to-be implies the following ought-to-do:

One ought to bring it about (*ceteris paribus*) that people respond to red objects in sunlight by uttering or being disposed to utter ‘this is red’ under appropriate promoting.

The ‘one’ here is vague, and making it more precise can help to distinguish between three possible views one might take about ‘ought-to-be’s. One relatively deflationary idea is that ‘ought-to-be’s materially imply a certain sort of conditional ‘ought-to-do’s. Generically,

(*) X ought to be φ

materially implies

If someone is responsible for X’s being φ , then that person ought to do what he/she can (*ceteris paribus*) to bring it about that X is φ .

⁴⁵ *Ibid.*

⁴⁶ I have changed Sellars’ example slightly to make it more intuitive and to make it more clearly fit the general formula stated above. He wrote, “(Ceteris paribus) one ought to respond to red objects in sunlight by uttering or being disposed to utter ‘this is red’” (*Ibid.*, p. 511).

⁴⁷ What about the linguistic rule “One ought not use unattached participles”, which seems true even if the person to whom it applies does not have the concept of an unattached participle? Sellars would insist that if this really is a true linguistic rule applying to all English speakers, we should view it as a rule of criticism (perhaps more precisely stated as: “One ought not to be disposed to use unattached participles”). And if it is rather meant as a rule of action, then it cannot apply to subjects who lack the relevant conceptual resources to follow it. The technical distinction between rules of action and rules of criticism is (in part) defined by whether the rule presupposes that the subject has the conceptual resources to follow it. (Thanks to John Broome for pressing me to make this point clearer.)

Call this the *conditional view* of the logical relationship between ‘ought-to-be’s and ‘ought-to-do’s.⁴⁸ A much stronger view sees ‘ought-to-be’s as materially implying universal ‘ought-to-do’s—generically, that is, (*) implies

Everyone ought to do what he/she can (*ceteris paribus*) to bring it about that X is ϕ .

Call this the *universal view* of the logical relationship between ‘ought-to-be’s and ‘ought-to-do’s. A third middle view sees ‘ought-to-be’s as materially implying nonconditional but also nonuniversal ‘ought-to-do’s—generically, that is, (*) implies

Someone ought to do what he/she can (*ceteris paribus*) to bring it about that X is ϕ .

Call this the *existential view* of the logical relationship between ‘ought-to-be’s and ‘ought-to-do’s. On this view, a true ought-to-be implies that someone is responsible for making the subject of the ought-to-be be the way it ought to be.

Sometimes Sellars seems to be adopting the universal view of linguistic rules of criticism, which seems to me to be the correct view of moral ‘ought-to-be’s, such as

People ought to feel outrage about genocide

which then be said to imply

Everyone ought to do what he/she can (*ceteris paribus*) to bring it about that people feel outrage about genocide.⁴⁹

However, Sellars’ leading example suggests the conditional view instead. This is

Clock chimes ought to strike on the quarter hour,

which seems to imply *at most* that

If someone is responsible for the (relevant) clock chimes, then that person ought to do whatever he/she can (*ceteris paribus*) to bring it about that they strike on the quarter hour.

This is made clearer by noticing the parallel to,

The bomb ought to spread shrapnel widely

which hopefully does not imply anything more than the conditional

If someone is responsible for the bomb’s spreading shrapnel widely, then that person ought to do whatever he/she can (*ceteris paribus*) to make the bomb spread shrapnel widely,

whose antecedent is, at least morally speaking, usually false. However, the intuitive idea of teaching our children things by doing what we can to get them to conform to particular rules of criticism suggests to me neither a universal ought-to-do nor a merely conditional ought-to-do. For instance, if it is true that a child ought to be able to tie his shoes by age four, then it would seem also to be true that someone (his guardians, but not everyone) has an obligation to teach him to be this way. This suggests the existential view. That is to say that

This child ought to be able to tie his shoes by age four

implies

⁴⁸ While “conditional”, it would be wrong to think of the resulting ‘ought-to-do’s as therefore “hypothetical”. The sense in which they are conditional is not that they are conditional on the agent’s antecedent ends/desires, which is what hypotheticality requires. The conditional-unconditional distinction cuts across the hypothetical-categorical distinction.

⁴⁹ Sellars uses a similar example where I think intuitions are less clear: “One ought to feel sympathy for the bereaved.” He suggests that this materially implies that one ought (*ceteris paribus*) to bring it about that people feel sympathy for the bereaved. See *Ibid.*, p. 509. It is important to remember that the *ceteris paribus* clause might be relied on quite heavily in cases where there are other obligations to mind one’s own business or to do something else more important than influencing people’s feelings about genocide.

Someone ought to do what he/she can (*ceteris paribus*) to bring it about that this child is able to tie his shoes by age four.

Which is the correct view of linguistic rules of criticism—the unconditional view suggested by the genocide example, the conditional view suggested by the clock example, or the existential view suggested by the shoe-tying example? I suspect that the latter view is correct in the linguistic case, but, more generally, I want to recognize the conceptual possibility of all three sorts of ‘ought-to-be’s. What is important to me here is that, in one way or another, Sellars’ idea of rules of criticism provides a good model for doxastic oughts. The idea would be to treat doxastic oughts as a species of rules of criticism (one of perhaps several distinctively epistemic species). They are of the form

X ought to be in doxastic attitude A towards proposition p under conditions C.

For instance,

X ought to believe that X is reading this paper right now under the condition that X is reading this paper right now in normal circumstances.

The crucial thing is that just like the other rules of criticism I have mentioned, viz.

The clock’s chimes ought to strike at the quarter hour,

All the beds ought to be made every morning,

A child ought to be able to tie his/her shoes by age four,

People ought to feel outrage about genocide,

and

One ought to respond to red objects in sunlight by uttering or being disposed to utter ‘this is red’ when appropriately prompted,

doxastic oughts seem to be true, so thinking of doxastic oughts on this model—whether it is as implying conditional, unconditional, or existential ‘ought-to-do’s—respects the true doxastic oughts principle. But doxastic oughts, conceived of as rules of criticism, are also consistent with doxastic involuntarism since, as the examples of the clock and the beds clearly reveal, rules of criticism do not have as a precondition on their truth that *their subjects* be capable of voluntarily following the rule. So thinking of doxastic oughts on this model also respects the no rewards principle. However, as Sellars’ account of linguistic rules of criticism reveals, none of this implies that believers cannot be agents. We just have to appreciate that they do not exercise agency in believe what they believe.

By respecting both the true doxastic oughts principle and the no rewards principle, this account avoids the drastic choice between doxastic involuntarism and epistemic deontology with which I began. Moreover, I think it surpasses the accounts suggested by Feldman and Kornblith. Let me explain.

Regarding Feldman: Even though he does not make this point, presumably some role oughts are rules of criticism—e.g. “Teachers ought to be interested in their subject.” However, such role-driven rules of criticism are not plausibly thought to be categorical. Good performance of the role of a kleptomaniac may require that one *be* cunning, but it is not a categorical rule of criticism that one ought to be cunning if one is a kleptomaniac. Rather, we want to say that, if one is a kleptomaniac, one ought to be treated by a psychiatrist. But, while this is perhaps a categorical rule of criticism, it is not a role-ought. And, likewise, doxastic oughts conceived of as rules of criticism seem to be categorical in a way that role oughts are not. So by treating doxastic oughts as rules of criticism that are not role oughts, my account gains important resources for explaining the apparent categoricity of doxastic oughts that Feldman’s account lacks.

Regarding Kornblith: Doxastic oughts do seem to me to derive from epistemic ideals, but not on the model of the derivation of typical *moral* ‘ought-to-do’s from *moral* ideals. This is because the lack of voluntary control we have over our beliefs is not the same sort of voluntary

control lacked by someone who, because she falls short of a moral ideal, is incapable of acting in the morally required ways. Even when we are epistemically ideal, we appear to lack direct voluntary control over our doxastic attitudes. So the possibility of being nonideal does not explain why the no rewards principle is true. By contrast, seeing doxastic oughts as epistemic rules of criticism does explain why the no rewards principle is true: Doxastic oughts are ‘ought-to-be’s and we cannot directly decide *to be* this way or that. In this way, then, the epistemic ideals which underwrite doxastic oughts are analogous to moral ideals about appropriate feelings. For example, perhaps it follows from the moral ideal of generosity that one ought to feel pleasure when giving to charity. If so, this can be true even though one does not have direct voluntary control over whether one feels pleasure when giving to charity. Because one lacks direct voluntary control, it does not make sense to offer to reward someone to have this feeling just like it does not make sense to offer to reward someone for having a particular belief. Though, of course, we can sometimes decide to do something that makes us be this way or that, and it can sometimes make sense to reward someone to do perform one of these actions with indirect influence over our states of feeling or belief. None of this is inconsistent with Kornblith’s account of doxastic oughts as deriving from epistemic ideals, but it moves beyond that account by providing a way to understand the difference between oughts that apply to actions and oughts that apply to beliefs.⁵⁰ The former are rules of action, while the latter are rules of criticism.

So my general idea is to treat doxastic oughts as rules of criticism that imply rules of action. Because rules of criticism are oughts which do not presuppose that their subjects have voluntary control over what is prescribed (or even that they can act at all), this explains why the no rewards principle is true, but since there are many clear examples of true rules of criticism, this helps to underwrite the true doxastic oughts principle. Now, although the connection I have claimed between doxastic oughts and rules of action rests heavily on the *ceteris paribus* clause and the notion of material implication (as it does for Sellars’ linguistic rules), I have claimed that doxastic rules of criticism imply rules of action. So it is fair to wonder: what rules of action are implied? I think this is a complex and difficult question that I do not hope to settle with any serious precision here. But I will conclude by gesturing at the kind of specifics I think are plausible and warning against one possible way of going that may seem to render the account trivial.

5. CONCLUSION: TOWARDS A FULLER ACCOUNT

There is nothing about thinking of doxastic oughts as ‘ought-to-be’s that precludes their being categorical. Here Kornblith’s idea that epistemic oughts derive from epistemic ideals is helpful.⁵¹ For example, it is plausible to think that people ought to feel outrage about genocide. With the distinction between rules of action and rules of criticism, this can be construed as a categorical rule of criticism deriving from a moral ideal of compassion and respect for persons. To be sure, it requires much reflection to identify epistemic ideals precisely. Much like our moral ideals

⁵⁰ It also makes sense of D. Owens’ claim in *Reason without Freedom: The Problem of Epistemic Normativity* (London: Routledge, 2000), p. 126, that “Some norms are not there to guide action, to govern the exercise of control: their function is to assess what we are.” Because he thinks this, Owens argues that not all responsibility presupposes freedom to choose—in particular, epistemic responsibility does not presuppose voluntary choice of our doxastic attitudes. I agree with this. However, I would say that epistemic responsibility is intimately related to responsibility to do things, over which we do have voluntary control. The conceptual connection between rules of criticism and rules of action that I have been urging allows one to make this connection.

⁵¹ For similar ideas compare C. Hookway, “Cognitive Virtues and Epistemic Evaluations,” *International Journal of Philosophical Studies*, 2 (1994), p. 225; and L. Zagzebski, *Virtues of the Mind* (Cambridge: Cambridge University Press, 1996), p. 219-31.

come from our natures as social beings, I suspect that our epistemic ideals will come from our natures as information tracking and transmitting beings, which is tied up with our natures as social beings. It is part of instituting and maintaining the social structures in which these ideals are ideals that we evaluate and react not only to each other's actions but also to each other's ways of being. These evaluations and reactions help to shape the way we are and, since the way we are is intimately tied to the way we will act, also how we will act. In such a situation, I think there is good reason to suppose that the doxastic oughts that derive from the ideals of being good information tracking and transmitting beings will be categorical rules of criticism that imply inter-personal epistemic rules of action.

There is one way to capture the categoricity of doxastic oughts which might seem to render the account I have sketched above trivial (even if correct). This can be brought into view by noticing that, although believing is a mental state and so not in the relevant sense an action, there is a very closely related mental act. This is the act of judging. This connection is highlighted and exploited by Shah and Velleman⁵² in their attempt to make sense of doxastic deliberation—that is, deliberation framed by the question whether or not to believe that *p*. They want to know how such deliberation can make sense from the perspective of doxastic involuntarism; and their suggestion is that doxastic deliberation is in the first instance deliberation about the mental act of judging. This they see as a species of the mental act of affirming a proposition. More specifically, it is the species of affirming a proposition, which is conceived of as subject to a particular norm of correctness, namely, the norm that the act is correct if and only if the proposition is true. In their view, this norm distinguishes judgment from other species of affirming, such as affirming a proposition for heuristic or polemical purposes, as a correlative norm of correctness distinguishes belief from other cognitive mental states, such as assuming and imagining. Shah and Velleman think this helps to avoid any problem doxastic involuntarism might be thought to cause for a coherent notion of doxastic deliberation. They write,

Now, there can be no problem about the possibility of deliberating whether to perform the mental act of affirming that *p*. As an *act*, mental affirmation is clearly eligible to be an object of deliberation. A problem might be thought to arise in the transition from that act of affirmation to an affirmative attitude, but to our minds, there is no problem about that transition, either. Exactly how one accomplishes the transition is of course ineffable, but it is a perfectly familiar accomplishment, in which a proposition is occurrently presented as true in such a way as to stick in the mind, lastingly so represented.⁵³

So, in the idiom of rules of criticism and rules of action, the idea would be to treat doxastic oughts as 'ought-to-be's which imply 'ought-to-do's about which acts of judging the agent should perform. So, for example,

You ought to believe that you are now reading this paper
would be a rule of criticism implying the rule of action

You ought to judge that you are now reading this paper.

Although they do not stress this point, one potential advantage of Shah and Velleman's view is that it can capture a sort of categoricity had by doxastic oughts that is not captured by the model of role oughts. This is because, in their view, it is part of the concept *judging* that an act's being a judgment consists in part in its being subject to the specific norms of correctness relevant to judgments. By contrast, it is not part of the *concept* of, e.g., being a slave that one is subject to any particular norm of correctness, such as that one ought to serve without question. To see this, notice that one who does not aim to serve without question may still count as a

⁵² *Op. cit.*

⁵³ *Ibid.*, p. 503.

slave, whereas one who does not aim at the truth in affirming a proposition does not count as judging. This means that, insofar as one is judging, one is subject to the oughts that derive from the constitutive norms of correctness for judgment. The same is not generally true of role oughts; it is not the case that, insofar as one is a slave, one is subject to the constitutive norms of being a slave.

Although I think this line of thought provides a novel way to capture the categoricity of doxastic oughts, the account of doxastic oughts is too simple for the purposes of coming up with a satisfying account of doxastic oughts in light of the apparent tension between the no rewards principle and the true doxastic oughts principle. This is for two reasons. First, even if judging is a mental act, it is not clearly (or at least always) a voluntary action. The no rewards principle applies to acts of judging as much as it applies to states of believing. Offered a high reward to *judge* that the U.S. is still a colony of Britain, you cannot just do it and collect the reward. So, treating these acts as the place where deontological evaluation can get a proper foothold in epistemology will provide no respite from the no rewards principle. Second, we have already seen that an important difference between rules of criticism and rules of action is that the latter but not the former have as a precondition on their truth that the subject to whom they apply has the conceptual resources needed to follow the rule. Part of the theoretical value of rules of criticism is that they apply regardless of whether the subject to which they apply has the conceptual resources necessary to follow them (e.g. “The clock chimes ought to strike at the quarter hour.”) If, however, for every true doxastic ought of the form

S ought to believe p,

we could derive a rule of action of the form

S ought to perform the mental act of judging that p,

then it would be incorrect to treat people as subject to doxastic oughts unless they had the concepts of *mental act* and *judging*. Moreover, if Shah and Velleman are right about what is involved in the concept of *judging*, one would also have to have the concept of *being subject to a norm of correctness*. But this does not seem right. Some of the true doxastic oughts apply to believers who do not themselves (yet) have the necessary conceptual apparatus to follow the ostensibly correlative rule of action about the mental act of judging. Small children and even some animals ought to believe certain propositions even if they do not (yet) have the conceptual apparatus needed to follow a rule of action about the mental act of judging.

In the theoretical milieu of a debate between doxastic involuntarists and proponents of epistemic deontology, the principal advantage of the distinction between rules of action and rules of criticism is that it shows how there can be a species of rule, properly so called, which does not require its subject, no matter who or what it is, to be capable of voluntarily following it. So I think it we would be giving up on this theoretical advantage if we thought that every doxastic rule of criticism implied a correlative rule of action prescribing the action of judging.

This is partially why, in my view, we need a more inter-personal conception of the relationship between doxastic oughts and the implied rules of action. One way to get this is to think of doxastic oughts as including a forward-looking element that is akin to the Sellarsian conception of language instruction. For example, if you have a doxastic attitude that you ought not to have, I think it is plausible to suppose that members of your epistemic community ought (*ceteris paribus*) to do what is in their power to disabuse you of this doxastic attitude by, for example, providing you with counter-evidence, counter-arguments, and, at the extreme, institutional care. (Again, due to the *ceteris paribus* clause, this rule of action is often void in the very many cases where there are more important things to do). However, doxastic oughts can also be plausibly thought to include a backward-looking element. For example, when we say,

You ought to disbelieve that the earth is flat,

it is plausible to suppose that this also implies both intra-personal rules of action such as

You ought to have read your science books and listened to your parents and teachers and inter-personal rules of action such as

Your parents and teachers ought to have taught you that the earth is not flat.

To go this way is to interpret doxastic oughts as logically related to 'ought-to-do's on the existential view discussed above. If a doxastic ought is true, then there is someone who ought do to (or ought to have done) something. Here, I have suggested vaguely that it is the believer's epistemic community (which might include his/her previous self) who is subject to this ought-to-do. I'm inclined towards this interpretation, though it does raise the difficult question of why we should think people in epistemic communities are responsible for each other's beliefs. I do not know of a completely satisfactory way to approach this question. However, it is important to notice that for my argumentative purposes in this paper, we could just as well go the other way and conceive of doxastic oughts as related to 'ought-to-do's on the conditional model discussed above. Perhaps doxastic oughts merely imply that, if someone is responsible for a particular believer's beliefs about something, then that person ought to do what he/she can to bring it about that the believer believes correctly. This would still be a case of treating doxastic oughts as rules of criticism. And like other rules of criticism, the key point is that the subject to which they apply does not have to have direct voluntary control over the state which is prescribed in order for the ought to be true. This is why the no rewards principle holds even in the face of the true doxastic oughts principle. Also, like other rules of criticism, doxastic oughts imply rules of action, and it seems plausible to see them as categorical rather than hypothetical and as implying both forward-looking and backward-looking inter-personal rules of action.