

TABLE OF CONTENTS

Patrick BONDY, Introduction.....	397
Sharon RYAN, In Defense of Moral Evidentialism.....	405
Andrew REISNER and Joseph VAN WEELDEN, Moral Reasons for Moral Beliefs: A Puzzle for Moral Testimony Pessimism.....	429
Dustin OLSON, A Case for Epistemic Agency.....	449
Benjamin WALD, Transparency and Reasons for Belief.....	475
Brian HEDDEN, Believing and Acting: Voluntary Control and the Pragmatic Theory of Belief.....	497
Notes on the Contributors.....	515
<i>Logos and Episteme</i> : Aims and Scope.....	517
Notes to Contributors.....	519

INTRODUCTION¹

Patrick BONDY

The ethics of belief, broadly speaking, has to do with how we ought to form, sustain, and revise beliefs. Philosophers have of course long been in the business of articulating rules for belief-formation, and there are a variety of questions to address and strategies for addressing them. Contemporary work falling under the label ‘the ethics of belief’ tends to draw much of its inspiration and guiding questions from the well-known exchange between W.K. Clifford and William James. In his essay, “The Ethics of Belief,”² Clifford argued for the strong evidentialist principle that it is always morally wrong to believe on insufficient evidence. In support of this principle, he described a case where a ship-owner acquires good evidence for thinking that his ship might not survive another voyage, but because he doesn’t want to spend money on repairs, he ignores the evidence, convincing himself that the ship will be fine. Later, the ship heads out to sea and sinks, killing everyone on board. Clifford argued that the ship-owner is clearly morally blameworthy for having held the evidentially unjustified belief that the ship could take another voyage. And, Clifford thought, the ship-owner would have been just as morally blameworthy for holding that belief even if the ship had managed to survive another trip and nobody had died.

In “The Will to Believe,”³ James famously replied to Clifford that there are cases where a choice about whether or not to believe *p* is *forced* (the choice cannot be avoided), *momentous* (it is important, and it has perhaps irreversible consequences), and *live* (both believing and refraining from believing are possible for the subject – in particular, the truth or falsity of *p* is not decided by the available evidence). James called these kinds of cases *genuine* options. When it

¹ The impetus for this special issue came from a workshop on the ethics of belief that was held at Cornell University, November 15-16, 2014, where earlier drafts of the papers by Sharon Ryan and Dustin Olson were presented. Special thanks are due to the Sage School of Philosophy for hosting the workshop, and to the participants and audience members who came out to the event. Thanks are also due the editors of *Logos & Episteme*, especially Eugen Huzum, for supporting and providing advice regarding this special issue.

² William K. Clifford, “The Ethics of Belief,” in *The Ethics of Belief and Other Essays* (Amherst, NY: Prometheus, 1999), 70-96. Originally published in 1877.

³ William James, “The Will to Believe,” in his *Essays in Pragmatism* (New York: Hafner, 1949), 88-109. Originally published in 1896.

comes to a genuine option, James argued, we are not rationally required to suspend judgment about p , as evidentialists would have us do. Because we cannot avoid choosing, and the evidence is insufficient to settle what to believe, we may allow non-evidential considerations to play a role in belief-formation in these kinds of cases.

We can tease out a number of interesting and important questions about the principles that Clifford and James put forward, and the picture of the mind and mental states that underpin the debate. For example, do Clifford and James presuppose that we have any serious sort of voluntary control over our belief-formation in claiming that there is a moral responsibility to believe what the evidence supports, or in claiming that there are cases where we may legitimately decide what to believe? Should the responsibility to believe what the evidence supports be understood as a moral responsibility, or is it more properly cast as an epistemic, or intellectual, responsibility (or is it both)? What is an acceptable level of evidential support for forming beliefs? Does every body of evidence support only one rational degree of confidence in any given proposition? Are there legitimate non-evidential reasons for belief? Would such reasons necessarily be pragmatic or practical reasons, or could there be non-evidential but still epistemic or intellectual reasons? And just what sorts of attitudes are beliefs, anyway?

This introduction is not meant to serve as a survey of work in the field,⁴ but it's worth pointing out that the ethics of belief as a sub-field of philosophy is alive and well.⁵ There are many defenses and criticisms of forms of evidentialism in the

⁴ For a more comprehensive overview, see Andrew Chignell's entry "The Ethics of Belief," in *Stanford Encyclopedia of Philosophy* (Spring 2013 Edition), ed. Edward N. Zalta, URL = <<http://plato.stanford.edu/archives/spr2013/entries/ethics-belief/>>.

⁵ Sometimes philosophers say things that can lead us to think that the ethics of belief in general as an area of research is under threat and needs to be defended, but these tend to be misleading statements. For example, Quine famously claimed that epistemology should be subsumed by empirical psychology, and that normative prescriptions for belief-formation should be replaced by the empirical study of how beliefs are actually formed in response to stimuli (see W. V. Quine, "Epistemology Naturalized," in his *Ontological Relativity and Other Essays* (New York: Columbia University Press, 1969), 69–90). But Quine later clarified that he wanted to retain a place for normative evaluations of belief. It's just that, in his view, the right kind of normative talk for epistemologists to engage in is instrumental in character: we should proceed by identifying the relevant cognitive goals, such as the achievement of true beliefs, and then we should proceed to identify good and bad ways to achieve those goals, and recommend the good ones. (See W. V. Quine, "Reply to White," in *The Philosophy of W. V. Quine*, ed. Lewis Edwin Hahn and Paul Arthur Schilpp (LaSalle, Illinois: Open Court, 1986), 663–665).

Another potentially misleading discussion of the ethics of belief can be found in Brian Huss, "Three Challenges (and Three Replies) to the Ethics of Belief," *Synthese* 168 (2009): 249–

literature,⁶ as well as arguments over instrumental conceptions of epistemic reasons and rationality,⁷ and arguments over whether we have any kind of control over our beliefs and what kind of control would be required for deontological terms of appraisal to be properly applicable to us as believers,⁸ and arguments over what kinds of things count as evidence at all.⁹

271. Huss identifies and replies to what he calls three challenges to the ethics of belief. The challenges Huss identifies are real, but they are challenges to particular views people have defended about the normative requirements on beliefs, not challenges to the business of working on normative requirements for beliefs in general.

⁶ Just a small sampling: Earl Conee and Richard Feldman, *Evidentialism: Essays in Epistemology* (Oxford: Oxford University Press, 2004), Nishi Shah, "A New Argument for Evidentialism," *The Philosophical Quarterly* 56, 225 (2006): 481-498, Pamela Hieronymi, "The Wrong Kind of Reason," *The Journal of Philosophy* 102, 9 (2005): 437-457, Andrew Reisner, "The Possibility of Pragmatic Reasons for Belief and the Wrong Kind of Reasons Problem," *Philosophical Studies* 145, 2 (2009): 257-272, Andrew Reisner, "A Short Refutation of Strict Normative Evidentialism," *Inquiry: An Interdisciplinary Journal of Philosophy* 58, 5 (2015): 477-485, and Trent Dougherty, ed., *Evidentialism and its Discontents* (Oxford: Oxford University Press, 2011).

⁷ See Richard Foley, *The Theory of Epistemic Rationality* (Cambridge, Mass: Harvard University Press, 1987), Richard Foley, *Working Without a Net: A Study of Egocentric Rationality* (New York and Oxford: Oxford University Press, 1993), Thomas Kelly, "Epistemic Rationality as Instrumental Rationality: A Critique," *Philosophy and Phenomenological Research* 66, 3 (2003): 612-640, Adam Leite, "Epistemic Instrumentalism and Reasons for Belief: A Reply to Tom Kelly's Epistemic Rationality as Instrumental Rationality: A Critique," *Philosophy and Phenomenological Research* 75, 2 (2007): 456-464, Thomas Kelly, "Evidence and Normativity: Reply to Leite," *Philosophy and Phenomenological Research* 75, 2 (2007): 465-474, and Clayton Littlejohn, *Justification and the Truth-Connection* (Cambridge: Cambridge University Press, 2012).

⁸ See William Alston, "The Deontological Conception of Epistemic Justification," *Philosophical Perspectives* 2 (1988): 257-299 (reprinted in his *Epistemic Justification: Essays in the Theory of Knowledge* (Ithaca: Cornell University Press, 1989), 81-114), Bernard Williams, "Deciding to Believe," in his *Problems of the Self: Philosophical Papers 1956-1972* (London: Cambridge University Press, 1973), 136-151, Barbara Winters, "Believing at Will," *The Journal of Philosophy* 76, 5 (1978): 243-256, Sharon Ryan, "Doxastic Compatibilism and the Ethics of Belief," *Philosophical Studies* 114 (2003): 47-79, Richard Feldman, "The Ethics of Belief," *Philosophy and Phenomenological Research* 60 (2000): 667-695, Richard Feldman, "Modest Deontologism in Epistemology," *Synthese* 161 (2008): 339-355, Rik Peels, "Believing at Will is Possible," *Australasian Journal of Philosophy* 88 (2014): 1-18, and Patrick Bondy, "Epistemic Deontologism and Strong Doxastic Voluntarism: A Defense," *Dialogue* (2015), DOI: <http://dx.doi.org/10.1017/S0012217315000487>.

⁹ See Timothy Williamson, *Knowledge and its Limits* (Oxford: Oxford University Press, 2000), John Turri, "The Ontology of Epistemic Reasons," *Nous* 43, 3 (2009): 490-512, Richard Feldman

The abundance of interest in and research on the ethics of belief is a very happy circumstance indeed, both because many of the questions falling under the heading of the ethics of belief are intrinsically interesting (to me, at least!), and because many of these questions are directly relevant to other sub-fields of philosophy and to the world more generally. For example: plausibly, there is a moral requirement to learn a reasonable amount about subjects that are morally important. (You morally ought to learn about what kinds of food babies can safely eat and what kinds will kill them before you feed your infant, for instance.) But there are different ways to take that requirement: are we required to gain lots of *knowledge* about subjects that are morally important? Or are we only required to do our best to gain knowledge (so that, for example, gaining justified but false beliefs is enough to meet the requirement)? What about if we gain justified true beliefs about the moral domain, but due to quirky features of the situation, we fail to have knowledge? And how much are we required to learn before we can stop and do other things? These seem like morally important questions to answer, and once we start addressing them, we're working on the ethics of belief.

A related area of research has to do with the epistemic norms associated with treating beliefs or propositions as reasons in practical deliberation. Epistemologists have done quite a bit of work on the epistemic norms of assertion,¹⁰ which is a special case of action, and some recent work has been done on epistemic requirements for treating propositions as reasons for acting more generally.¹¹ One particularly pressing question has to do with the relation of justifications, excuses, and control. For example: if Will is hosting a dinner party, and he serves a dish which contains an ingredient to which his guest Wanda is deathly allergic, we might naturally be inclined to morally blame Will for his oversight. It's normally expected that hosts inquire into the allergies of their

and Earl Connee, "Evidence," in *Epistemology: New Essays*, ed. Quentin Smith (Oxford: Oxford University Press, 2008), 83-104, and Littlejohn, *Justification and the Truth-Connection*.

¹⁰ e.g. Timothy Williamson, *Knowledge and Its Limits*; Rachel McKinnon, "The Supportive Reasons Norm of Assertion," *American Philosophical Quarterly* 50, 2 (2013): 121-135, John Turri, "Knowledge and Suberogatory Assertion," *Philosophical Studies* 167, 3 (2014): 557-567, Jonathan Kvanvig, "Assertion, Knowledge, and Lotteries," in *Williamson on Knowledge*, ed. Duncan Pritchard and Patrick Greenough (Oxford: Oxford University Press, 2009), 140-160, B. J. C. Madison, "Is Justification Knowledge?" *Journal of Philosophical Research* 35 (2010): 173-191, and Jennifer Lackey, "Norms of Assertion," *Nous* 41, 4 (2007): 594-626.

¹¹ e.g. John Hawthorne and Jason Stanley, "Knowledge and Action," *Journal of Philosophy* 105, 10 (2008): 571-590, Mikkel Gerken, "Warrant and Action," *Synthese* 178, 3 (2011): 529-547. Susanne Mantel, "Acting for Reasons, Apt Action, and Knowledge," *Synthese* 190, 17 (2013): 3685-3888, and Clayton Littlejohn, *Justification and the Truth-Connection*.

guests, after all. But if we later find out that Will did inquire about allergies, and Wanda didn't reply because she didn't even know that she had the allergy in question, then we would of course withdraw our blame for Will. But beyond the appropriateness of withdrawing our blame, it's not entirely clear how we ought to regard Will and his dinner. Should we view Will as having done exactly what he ought to have done, since he did his best to serve food everyone could eat? Or should we only think that Will is *excused* or *blameless* because he couldn't know about Wanda's allergy, but that he was nevertheless *unjustified* in serving the dinner he did? After all, he didn't know that his meal was safe for his guests to eat. How we address these questions will depend, among other things, on what status beliefs must have in order to be properly treated as premises in practical deliberation. This is an area of research that deserves further development.

This special issue collects five new essays on various topics relevant to the ethics of belief. The issue begins with Sharon Ryan's paper, "Moral Evidentialism," in which Ryan defends Clifford's evidentialist principle. There are of course a number of recent epistemologists who defend evidentialist principles, but they typically reject Clifford's principle that it's *morally* wrong to believe on insufficient evidence. The standard evidentialist view tends to be that it's *epistemically* wrong, or impermissible, or unjustified, to form beliefs on insufficient evidence, but that Clifford went too far with his moral condemnation of evidentially unsupported beliefs. One common objection to Clifford's view is that it just seems too strong: there seem to be cases where it can be morally permissible and pragmatically justified to hold beliefs that go against the evidence.¹² Another kind of objection is that we need to have voluntary control over anything for which we can be morally responsible, and that we don't seem to have voluntary control over what we believe.¹³

Ryan argues that there are convincing replies to both of these sorts of objection. For example, against the second objection, she argues that we don't after all need to have voluntary control over events or actions in order to be praiseworthy or blameworthy for them. And furthermore, she argues, it seems that if we are going to make sense of the practice of holding people morally responsible for their actions, then we need to be able to hold people morally responsible for their beliefs, when their beliefs are held contrary to good evidence.

¹² John Heil gives a standard case meant to illustrate this possibility, in "Believing Reasonably," *Noûs* 26, 1 (1992): 47-62.

¹³ Alston, "Deontological Conception," gives an oft-cited formulation of this argument against doxastic obligations.

So the standard rejection of a moral version of evidentialism deserves at the very least to be seriously reconsidered.

The second paper in this issue is Andrew Reisner and Joseph Van Weelden's (RVW) "Moral Reasons for Moral Beliefs: A Puzzle Case for Moral Testimony Pessimism." *Moral testimony* is testimony to the effect that some moral claim is true, or that some action is morally right, and a *moral expert* is someone who is more likely than a non-expert to arrive at true beliefs about moral matters. Moral testimony *optimists* think that when an identifiable moral expert has given moral testimony, non-experts may legitimately form moral beliefs on the basis of that testimony, other things being equal (e.g. other identifiable moral experts mustn't be known to have given conflicting moral testimony). Moral testimony *pessimists* deny the legitimacy of forming moral beliefs in that way.

RVW construct a problem case for moral testimony pessimists. The case is designed to show that it is very natural to think that non-experts may legitimately form beliefs on the basis of the testimony of moral experts, because when non-experts form beliefs in this way, this makes it more likely that they will form true moral beliefs, and consequently that they will perform morally right actions. Requiring that we never form moral beliefs on the basis of expert testimony makes it likelier that we will perform morally wrong actions. RVW go on to consider a number of arguments for pessimism, such as Alison Hills's argument that it is better to have moral understanding than to have moral knowledge without understanding (where it's possible to gain moral knowledge but not moral understanding by forming beliefs on the basis of the testimony of moral experts).¹⁴ They concede that this and other related arguments may be enough to show that there is some *pro tanto* reason not to form beliefs on the basis of moral testimony, but they argue that there is always also a contrary reason in favour of optimism. And, RVW argue, there doesn't seem to be any reason to think that the *pro tanto* reason in favour of pessimism ever decisively favours pessimism in any concrete case.

In "A Case for Epistemic Agency," Dustin Olson describes the concept of epistemic agency, and argues that it has a place in our theorizing about the formation and justification of beliefs. According to Olson, epistemic agency is a kind of agency which we exercise over our belief-formation. He argues that belief-formation is a skill, and like any other skill, it can be developed and refined. And, because we can improve or fail to improve our belief-forming skills, we can be better or worse at forming beliefs in various domains, and so normative evaluations of the way we form our beliefs can be appropriate.

¹⁴ Alison Hills, "Moral Testimony and Moral Epistemology," *Ethics* 120, 1 (2009): 94-127.

Central to Olson's account is that epistemic agency does not presuppose either direct or indirect doxastic voluntarism: you can exercise epistemic agency with respect to your formation of the belief that p even without ever having anything like an intention to form the belief that p . Olson goes on to defend the use of the concept of epistemic agency against the challenge that the mechanistic character of belief-formation rules out the possibility of epistemic agency,¹⁵ and that employing a concept of epistemic agency doesn't respect the sense in which it's not possible to practice epistemic self-improvement.¹⁶

In "Transparency and Reasons for Belief," Benjamin Wald considers the relation between the aim of belief and the transparency of doxastic deliberation. It is commonly held among epistemologists that belief aims at the truth, although there are several importantly different ways to understand what this aim-of-belief-talk amounts to. According to 'normativists' about the aim of belief, it is partly constitutive of the mental state of belief that any belief is *correct* if and only if it is true (where 'correct' is supposed to be more than just a synonym for 'true'). And, according to Shah and Velleman,¹⁷ appealing to the aim of belief in this normative sense can help us to explain what they call the 'transparency' of doxastic deliberation, which is the fact that in consciously deliberating about whether to believe p , we automatically deliberate directly about whether p is true, rather than, say, about whether it would be a good thing to believe that p . If belief *constitutively* has a (normative) truth-aim, the explanation goes, then anyone conceptually sophisticated enough to deliberate about whether to believe p must *already* endorse an evidential norm on belief. So the normative truth-aim of belief explains why we can only appeal to evidence for or against the truth of p in deliberating about whether to believe p .

Wald argues that Shah and Velleman's explanatory strategy fails. He agrees that it seems to be a conceptual truth that beliefs cannot be deliberately held on the basis of non-evidential reasons, but he argues that if anyone were to form a belief on the basis of such a reason, they would not thereby be rationally criticisable. But Shah and Velleman's aim-of-belief explanation of transparency entails that if anyone were to form a belief on the basis of a non-evidential consideration, they *would* thereby be rationally criticisable. So their explanation seems to be mistaken. Wald's positive strategy to explain why only evidential

¹⁵ See Hilary Kornblith, *On Reflection*, (Oxford: Oxford University Press, 2012).

¹⁶ See Kristoffer Ahlstrom-Vij, "Why We Cannot Rely on Ourselves for Epistemic Improvement," *Philosophical Issues* 23 (2013): 276-96.

¹⁷ See Nishi Shah and David Velleman, "Doxastic Deliberation," *The Philosophical Review* 114, 4 (2005): 497-534.

reasons seem to be genuine normative reasons for belief involves combining the view that transparency is conceptually true together with a moderate form of motivational internalism about epistemic reasons.

Finally, in “Believing and Acting: Voluntary Control and the Pragmatic Theory of Belief,” Brian Hedden explores an interesting and novel kind of control which we can exercise over our beliefs, if the pragmatic account of belief is correct. According to the pragmatic account, whether a subject *S* believes that *p* depends in part on how well an attribution of the belief that *p* to *S* would help render *S*’s actions rationally intelligible. The pragmatic account of belief is of course controversial, but it does enjoy a certain amount of intuitive support – for example, if *S* appears to sincerely claim to believe that *p*, but we are unable to make *S*’s actions seem at all rationally intelligible except by attributing to *S* the belief that *not-p*, then we might naturally be inclined to attribute to *S* the belief that *not-p*, and conclude that *S* is confused about the content of his beliefs.

If our actions constrain our beliefs in this way, Hedden argues, then it is possible to exercise voluntary control over our beliefs, in cases where our performing or failing to perform an action will be partly constitutive of our having or not having a particular belief. (After all, we do typically have voluntary control over our actions.) And, Hedden argues, this indirect sort of voluntary control over our beliefs might be sufficient to save a responsibilist conception of epistemic justification or evaluation from ought-implies-can objections of the kind given by Alston.¹⁸

The papers collected here address various themes from the ethics of belief. They shed fresh light on important questions, and bring new arguments to bear on familiar topics of concern to most epistemologists, and indeed, to anyone interested in normative requirements on beliefs either for their own sake or because of the way such requirements bear on other domains of inquiry.

¹⁸ Alston, “Deontological Conception.”

IN DEFENSE OF MORAL EVIDENTIALISM

Sharon RYAN

ABSTRACT: This paper is a defense of moral evidentialism, the view that we have a moral obligation to form the doxastic attitude that is best supported by our evidence. I will argue that two popular arguments against moral evidentialism are weak. I will also argue that our commitments to the moral evaluation of actions require us to take doxastic obligations seriously.

KEYWORDS: W.K. Clifford, ethics of belief, doxastic obligations, evidentialism, moral evidentialism

What people understand, or have the capacity to understand, is morally significant. Cognitive states and capabilities explain, in part, why we hold most adult human beings morally responsible for their actions. Human babies, in contrast, are not morally responsible agents because their cognitive abilities are still too undeveloped. Moral responsibility also seems to turn, in part, upon what we believe, know, or are expected to know about the consequences of what we are doing. If I know that my behavior is likely to cause an innocent person to suffer terribly and I do it anyway, I am much worse than someone who performs the same action, but with good reason to think that nobody will suffer. Knowingly causing unnecessary suffering is especially bad.¹ This all seems uncontroversial. But does it make sense to say that our beliefs themselves are open to moral appraisal? More specifically, do we have *moral obligations* about the doxastic attitudes we form? I will spend this paper defending moral evidentialism, the view that we have a moral obligation to form the doxastic attitude that is best supported by our evidence. I will argue that two popular arguments against moral evidentialism are weak. I will also argue that our commitments to the moral evaluation of actions require us to take doxastic obligations seriously.

Clifford's Radical Evidentialist Principle

If this sounds like an exhumation of W. K. Clifford, to some extent, it is! I think Clifford's view has been unfairly dismissed, and I will take some time to try to restore a fresh interest in his position. In his 1877 essay, "The Ethics of Belief," Clifford defended the view that we have a moral obligation to never believe

¹ For example, think about how our moral evaluation of Takata (the airbag maker) changed as it became apparent that the company knew their airbags were defective.

Sharon Ryan

anything on insufficient evidence. According to Clifford, "It is wrong always, everywhere, and for anyone to believe anything on insufficient evidence."² I will not be attempting a historical exegesis of Clifford here. Instead, I will focus on his general idea that we have moral obligations that pertain to our doxastic attitudes. I will begin this discussion by distinguishing a view I will call Clifford's Principle, thereby acknowledging that it has similarities to Clifford's thinking on the subject, and argue that a plausible interpretation of it offers important instruction and guidance for living a good life and cultivating wisdom.

Clifford supported his principle with a story of a shipowner who had concerns about the safety of his ship. After a thorough examination of the ship, a well-known and reliable safety inspector documented serious safety violations. The inspector recommended that the ship undergo extensive repairs before sailing. Because those repairs would be costly, the shipowner ignored the inspector's report, convinced himself that the ship could make one more journey, and confidently sent the ship off on what he had hoped to be another lucrative cruise. Unfortunately, the safety inspector was correct in his diagnosis, and the ship was not able to make one more voyage. It sunk to the bottom of the ocean, killing all of the passengers and crew on board. The shipowner is legally and morally responsible for the death of the passengers. Everyone grants that lesson from the story. The controversial lesson of the story is Clifford's insistence that the ship owner is also morally guilty for holding the belief that the ship was seaworthy. Clifford claims that whether or not the ship sunk, he was guilty for believing as he did. I will begin the discussion of whether Clifford is correct with the following statement of his view.

(CP1) It is always morally wrong to believe any proposition p on insufficient evidence.

For the purposes of this paper, I think we can gloss over what makes a body of evidence sufficiently strong to justify the attitude of belief. It might, however, be worth distinguishing this thesis from another, more informative thesis. CP1 only tells us when it is wrong to form a belief. Perhaps a more instructive thesis that tells us which of the three possible doxastic attitudes is morally right or wrong would be even more interesting to think about. Consider CP2, a variation of CP1:

(CP2) S *morally ought* to have doxastic attitude D (belief, disbelief, suspension of judgment) toward p at t iff having D fits S 's evidence at t .

² William K. Clifford, "The Ethics of Belief," in *The Ethics of Belief and Other Essays* (Amherst, NY: Prometheus, 1999), 77. Originally published in 1877.

CP2 tells us that whenever a doxastic attitude is epistemically justified, we have a moral obligation to hold that attitude. I believe Clifford would be satisfied with CP2 as an accurate representation of his view.

There are other ways one might spell out evidentialist positions on the ethics of belief. One such way is:

(CP3) *S epistemically ought* to have doxastic attitude *D* toward *p* iff believing *p* fits the evidence *S* has toward *p*.

CP3 is basically the well-known and widely discussed evidentialist thesis defended by Richard Feldman and Earl Conee.³ In their decades of work devoted to defending evidentialism, Feldman and Conee are describing purely *epistemic* obligations.⁴ That is, obligations about what it takes to believe rationally or justifiably (in the sense of a necessary condition for knowledge). Feldman and Conee are *not* claiming that we have a *moral* obligation to have epistemically justified beliefs. Thus, they would not endorse CP2. Although I think Clifford would find CP3 attractive, it fails to capture Clifford's signature, moral stance.

Another possible way of characterizing an evidentialist view about doxastic obligations is:

(CP4) *S prudentially ought* to have doxastic attitude *D* toward *p* iff believing *p* fits the evidence *S* has at *t*.

CP4, like CP3, does not claim that we are morally required to believe what our evidence supports. However, according to CP4, believing in accord with our evidence is what is practically obligatory. This is because believing in accordance with our evidence leads to practical advantages. Perhaps those who believe in accord with their evidence are more likely to find food and shelter than those who rely on wishful thinking or astrology. I am dubious that we have practical *obligations*. There are, for sure, practical costs and benefits of actions, and those costs and benefits can have some impact on the moral status of our actions. And, some actions are more practically advantageous than others, but I am not sure that there is a purely practical sense of 'obligation.'

³ For their early articulation and defense of evidentialism, see Richard Feldman and Earl Conee, "Evidentialism," *Philosophical Studies* 48, 1 (1985): 15-34. See their *Evidentialism: Essays in Epistemology* (Oxford University Press, 2004) for a more recent and more developed discussion of evidentialism.

⁴ CP3 is actually a bit different from Feldman and Conee's view in that they use the term 'justified' rather than 'ought.' Feldman's work on the ethics of belief indicates that he is willing make use of the concept of an epistemic obligation in terms of believing what one's evidence supports. Nothing in this paper depends upon CP3 being the actual thesis Feldman and Conee endorse.

Sharon Ryan

Another possibility is to combine all of the alleged types of obligation noted and claim that believing what one's evidence supports is what is obligatory in all three ways. That is what (CP5) says:

(CP5) S morally, epistemically, and prudentially ought to have doxastic attitude D toward p iff believing p fits the evidence S has toward p .

And finally:

(CP6) S ought, all things considered, to have doxastic D toward p iff believing p fits the evidence S has toward p .

I am deeply perplexed by the idea that there is any real obligation captured by the notion of an all things considered obligation that is anything more than what is captured by CP4 or CP5. But that is what CP6 is claiming. Susan Haack⁵ and Richard Feldman⁶ have each provided interesting discussions of various senses of obligation at work in the ethics of belief literature. Here, I simply note a few ways of working out some evidentialist positions. Because I am not convinced that there are any epistemic obligations that are not subspecies of moral obligations; I do not think we have practical obligations, although I do acknowledge the idea of something being practically advantageous; and I do not really understand what an 'all things considered' obligation is, I will not be defending any of those ideas. Since I think CP2 is more interesting than CP1; extremely plausible on its face; strongly Cliffordian in spirit; less mind-boggling than CP3-CP6; and controversial enough for this paper, it is the version of Clifford's view I will focus upon. I am not insisting that this is the most accurate interpretation of Clifford's view. My interest in this paper is merely to try to articulate and defend what I take to be a promising position in the ethics of belief. From here on out, when I make reference to Clifford's Principle (CP), I have CP2 in mind. CP entails the view that we have moral obligations about what we believe. It also entails an evidentialist thesis specifying that those moral obligations are determined, exclusively, by our evidence.

Objection #1 to Clifford: The Inflexibility Problem

These days, CP is not taken seriously. Two lines of objection have been taken to show that Clifford's view is mistaken. One line of argument challenges Clifford's idea that doxastic obligations are determined by, and only by, our evidence. The

⁵ See Susan Haack, "The Ethics of Belief Reconsidered," in *The Philosophy of Roderick M. Chisholm*, ed. Lewis E. Hahn (LaSalle, IL: Open Court, 1997), 129-144.

⁶ See Richard Feldman, "Epistemic Obligations," *Philosophical Perspectives* 2 (1988): 235-255.

second line of argument challenges the idea that we have obligations, of any kind at all, about what we believe. According to the second objection, doxastic attitudes lack a feature that is necessary for a genuine moral obligation. I will address both sorts of challenge. I will argue that both challenges are much weaker than they have seemed to many people, and that's a good thing, since holding people responsible for their behavior depends, in part, on holding people responsible for what they believe.

I'll call the first challenge 'The Inflexibility Problem.' The Inflexibility Problem is the criticism that CP is too rigid in that it focuses exclusively on evidence in determining our doxastic obligations. The Inflexibility Problem is not an attempt to show that we do not have doxastic obligations. As I understand the Inflexibility Problem, it alleges that many of us find ourselves in situations where it is acceptable to have a doxastic attitude that is not supported by our evidence. In "The Will to Believe," William James, though quite sympathetic to evidentialism, rejected CP because it judges religious beliefs that are not supported by one's evidence to be immoral. James argued that in cases where belief is live, forced, and momentous, it is morally permissible to believe, even if we know that we lack sufficient evidence.⁷

In his *Stanford Encyclopedia of Philosophy* essay, "The Ethics of Belief," Andrew Chignell presents a non-theological case.

Suppose that you would like to retain a good relationship with your daughter, and you are aware that this requires believing the best of her whenever possible. You have some moderate but not compelling evidence for the proposition that she is using drugs in the house when you are away (in response to your queries, she claims that she has recently taken up meditation, and that the funny smell when you come home is just incense). Still, if your relationship will be seriously damaged by coming to view your daughter as a habitual drug-user, then you seem to violate a prudential norm if you believe that she is. In other words, it is prudent, given your ends to withhold belief about the source of the aroma altogether, or to believe, if possible, that she is burning incense in your absence. On the other hand, if you regard the occasional use of recreational drugs as harmless fun that expresses a healthy contempt for authority, then it might be prudent for you – confronted with the telltale odor – to form the belief that your daughter has indeed taken up the habit in question.⁸

⁷ William James, "The Will to Believe," in his *Essays in Pragmatism* (New York: Hafner, 1949), 88-109. Originally published in 1896.

⁸ Andrew Chignell, "The Ethics of Belief," in *Stanford Encyclopedia of Philosophy* (Spring 2013 Edition), ed. Edward N. Zalta, URL =<http://plato.stanford.edu/archives/spr2013/entries/ethics-belief/>, 5.

In Chignell's case, CP does not seem to accommodate our intuition that a parent's relationship to his or her child is more valuable, morally and prudentially, and ought to be given preference over believing what the evidence supports.

Another widely discussed example involves a woman discovering evidence that her husband is cheating on her. She finds lipstick on his collar, a piece of paper with another woman's phone number on it in his pocket, and the like. In this case, to preserve the marriage, it is allegedly morally and prudentially best for her to believe her husband is not cheating on her despite the evidence to the contrary.⁹ Again, CP is too rigid.

Another example involves a patient who is diagnosed with cancer. Most similarly diagnosed patients, suppose, regardless of their beliefs about their likelihood of survival, die of the disease within eight weeks. However, it has been shown that those who believe they will not die have a slightly better survival rate than those who believe what the evidence supports. Again, Clifford's recommendation is to believe what your evidence supports – always and everywhere. So, despite such an important, though unlikely, possible benefit, Clifford's principle commands that the cancer patient believe she will likely die.

Another case involves the value of confidence. When the Pirates' centerfielder, Andrew McCutchen (one of the best hitters in the major leagues), comes up to bat, should he believe he is going to get a hit, despite the realization that the odds are approximately 3-1 against him? Certainly! But, again, Clifford won't budge. "It is wrong always, everywhere, and for anyone to believe anything on insufficient evidence."¹⁰

Another popular example involves the moral demands of friendship. Suppose your friend, who claims to be innocent, is accused of a terrible crime and the evidence presented against her is very strong. If so, at this point in the investigation, the belief that your friend is guilty is epistemically justified for you. Despite that, it seems that you have a moral obligation, as a loyal friend, to trust your friend, and believe in her innocence, despite the evidence, until there is no possible room for doubt. And, even then, one might argue, one should believe in the innocence of one's friend. Again, CP cannot tolerate ignoring what the evidence supports.

Finally, what if believing against your evidence is the only way to save innocent lives? Wouldn't it be morally obligatory to violate Clifford's principle? Imagine that I am visiting a huge art museum and a very untalented, violent, and insecure artist (with a 100% reliable lie detector) asks me if I believe his work,

⁹ John Heil, "Believing What One Ought," *Journal of Philosophy* 80 (1983): 752-85.

¹⁰ Clifford, "The Ethics of Belief," 77.

which I am currently viewing, is brilliant. Suppose, based on all of my overwhelming visual evidence and superb aesthetic sensitivity, I believe his work is awful - not brilliant at all! What should I do? If I tell him I don't think his work is brilliant, he will blow up the art gallery and kill all of the people and destroy many incredible works of art. If I lie, he will detect the lie and he will blow up the museum anyway. Morally, it seems that despite my evidence, I must actually believe his work is brilliant. What does CP say about this? Well, it seems that CP demands that I believe his work is awful since that's what my evidence supports. Believing against your evidence is always morally wrong - no matter what the consequences.

Thus, Clifford's radical evidentialism has been abandoned because it seems too strict. It seems that there are occasions when purely epistemic considerations should not determine what one ought to believe.

Problems with the Inflexibility Problem: A Defense of Moral Evidentialism

Although all of the examples described are interesting, I don't think any of these examples show that we do not have a moral obligation to have epistemically justified beliefs.

James' example is unconvincing for several reasons. James' characterization of theistic belief as forced is a mistake. Therefore, on James' own criteria, this is not a situation in which one would be permitted to believe without sufficient evidence. One does not have to either believe God exists or believe God does not exist. One can always suspend judgment. And, if that is what one's evidence supports, that is what Clifford judges we ought to do. Moreover, it is implausible that for most people, the question of whether God exists is a question for which the evidence is, or seems to the inquirer to be, balanced or neutral. Many people think that the question of God's existence can be decided by reason and it is not one of the outlier cases that James allows to be decided by passion. The fact that theistic belief is so monumental for so many people, and the fact that it has been used to justify all sorts of morally significant behavior and policy, makes it all the more important to be extremely careful when forming one's beliefs. Moreover, it is not at all clear that *believing* in God delivers practical, intellectual, and moral payoffs that outweigh the practical, intellectual, and moral costs. Attitudes other than belief, given all that Clifford warns us about the costs of evidentially unsupported belief, would be preferable. Any benefits thought to come from belief can likely be obtained effectively by *hoping or wanting it to be true* (or perhaps even having faith if faith is understood as a psychological and emotional stance distinct from belief) that God exists and engaging in spiritual practices. Finally, a

Sharon Ryan

defender of moral evidentialism should respond by noting the obvious. Sometimes doing what's morally right is not practically, intellectually, or emotionally satisfying. Sometimes, morality demands sacrifice, and James has not shown that the payoffs of theistic belief yield the result that believing with insufficient evidence is *morally* permissible. At best, James has a challenge for CP4, CP5, or CP6. But this is certainly not a clear-cut and convincing objection to Clifford's view as articulated in CP2.

Chignell's example is a bit complicated, and it is probably intended to be an objection to a view such as CP4, CP5, or CP6. In that case, it is irrelevant to our concerns here and the criticisms I mention are unfair to Chignell's actual point. (Again, remember that when discussing Clifford's Principle in this paper, I am focusing on CP2.) Despite Chignell's actual intention, I think his example is interesting to consider as a potential objection to CP2, and I will take it as such here. What is especially attractive about Chignell's example is that it appears to be an ordinary example that actual people encounter in life quite frequently. If successful, it shows that Clifford is way off-track since moral demands often pull us away from having epistemically justified beliefs.

To be a genuine counterexample to CP, it must be understood as an example in which the parent morally ought to, for the sake of the relationship, believe the kid is not smoking pot in the house, but Clifford's principle demands that the parent believe she is smoking pot in the house (or depending on the evidence, that the parent should suspend judgment). It is not clear to me what the evidence supports in this example, so it is not clear to me what doxastic attitude Clifford's principle demands. Let's just stipulate that the evidence supports believing the kid is smoking pot in the house. Clifford's response, on this interpretation of the facts of the case, should be to question the empirical assumption that what is morally best for the relationship is for the parent to believe what the kid says and ignore the other relevant evidence. Doing so may be the easiest way to deal with the issue, but that does not mean that is the morally best approach. Intuitively, it seems that the best relationships are not built on lies, deception, and repression of the truth, but are built on dealing honestly and fairly with whatever version of reality is supported by the evidence. So, if interpreted as an objection to CP, and one that is representative of a type of situation that we encounter frequently, we should reject, or at least doubt, the idea that the parent must choose between CP and doing what is best for the relationship. Following CP, having a healthy respect for the evidence, and living with kindness, openness, and love sounds like the morally best way to deal with the situation. Clifford should argue that this is a case in which following CP's recommendation is what will lead to the strongest and

most loving relationship. And, even if it actually would somehow be morally best for the relationship to be mixed up with this lie, that does not show that it is morally best, overall, to believe against one's evidence. It might be morally best, overall, to focus on the long-term consequences for everyone involved and all that unravels as a result of the lie. That is, perhaps considering all of the morally relevant facts, risking the relationship is what is morally best. Of course, we could simply stipulate that, somehow or other, believing she is not smoking pot has enormous moral value that outweighs all the other morally valuable factors. I will consider a different case, later, that does a better job at making such a case. For now, I merely want to note that this example is not an ordinary, simple, clear-cut counterexample to moral evidentialism.

The case of the cheating husband is not convincing either. To begin, if all it is supposed to show is that it may be imprudent to do the morally right thing, that's consistent with CP (again understood at CP2). To refute CP, this should be a case where one is epistemically justified in believing p (my husband is cheating on me), but she is *morally obligated* to disbelieve p (or suspend judgment on p .) But, it is very difficult to work out the details so that the case is convincing. If the marriage is worth preserving, and it is not obvious from the sparse details that it is, normally the best way of salvaging a worthwhile relationship is to accept the facts and deal with them in an open and honest way. How is a relationship based on lies and deception better than one based on openness and honesty? Again, I do not think that Clifford should concede that there is, in fact, a genuine conflict here. Doing what's morally best for a relationship is consistent with, and in fact ordinarily requires, following one's evidence. Furthermore, even if it could somehow or other be shown that the relationship cannot endure the truth, and yet the relationship is of great moral value, this does not show that it is morally best to preserve the relationship rather than accept the truth. Perhaps the other woman's life (the one whose lipstick is on the husband's collar) would benefit from the truth being known. Or, perhaps accepting the truth would lead to better outcomes for everyone else impacted by the truth. So, this example is not a clear and convincing refutation of Clifford's Principle.

The cancer case is better, but not convincing if it is intended to be a realistic problem that should make us dubious of Clifford's basic idea. If this is to be a realistic scenario, the cancer patient should remain hopeful and positive, but that's possible without going to the extreme of actually believing, against her evidence, that she is going to live. On any realistic telling of the case, she should believe there's some small possibility that she will survive and she should do her very best to remain positive, do whatever she can to get as healthy as possible, and not give

Sharon Ryan

up on the fight. However, she should believe that it is very likely that she will die, and she should prepare herself and her loved ones for that likely scenario, while using a belief about the possibility of healing, and other positive and rational beliefs, to motivate her. Of course, she should not *dwell on* the negative belief that she has a high probability of dying, and that's not what Clifford recommends. He's not claiming, and nobody who is attracted to CP should claim that we have a moral obligation to dwell upon negative and depressing beliefs. All CP entails is that the cancer patient has a moral obligation to believe that it is likely that she will die within eight weeks. That's not cold, inflexible, irrational, or immoral. And, it is totally consistent with hoping, praying, feeling optimistic, or whatever, for a positive future.

The case of confidence is no better as an objection to moral evidentialism. If the example shows that some epistemically justified beliefs do not contribute to confidence, that's not a problem for Clifford's principle. Andrew McCutchen should be confident. After all, given his batting average and athletic prowess, he's got a better chance than almost anyone else of getting a hit. Believing he is a great baseball player and believing that he has a better chance than almost anyone of getting a hit are morally acceptable beliefs on Clifford's view. He is a great baseball player, and he has every reason to believe so and to feel confident. However, being rationally self-confident does not license him to believe he will get a hit. Believing he will get a hit would be irrationally arrogant even if it is helpful to his batting performance. If he's smart, and he follows the dictates of CP, he won't dwell on the rational belief that he is unlikely to get a hit. Actually, he probably shouldn't dwell on any beliefs at all when he's on the field and in the batter's box. He should spend all of his mental energy focusing on the ball!

The friendship case is tricky, but ultimately unconvincing. Sometimes, what we learn about our long-standing friends provides us with extremely good evidence about their moral character. Depending on how this example goes, the evidence we have about a friend's moral character can provide a powerful defeater to what would otherwise appear to be clear evidence of our friend's guilt. Without more details about what the evidence actually supports in this situation, it is unclear what CP yields as the morally justified doxastic attitude to hold. Assuming that the totality of the evidence, including all you know about your friend's character, really supports believing she is guilty, then I think that's what you morally ought to believe. That by no means rules out helping her find a good lawyer, making her a cake, visiting her in prison, and engaging in other forms of friendly support. It is difficult to imagine that a 'friendship' laced with deception

has greater moral value than a friendship built on honesty and acceptance. Accepting the facts, it seems to me, would allow you to be the best possible friend.

So far, I'm suggesting that the moral value of believing against one's evidence is actually much lower than these examples suggest. The awful art case is the most difficult case for Clifford since there is enormous moral value in believing against one's evidence. However, the cost is that this type of case is very unusual. Normally, our justified beliefs, all by themselves, do not have morally awful consequences. Normally, we can separate our beliefs from our behavior. (For example, you should believe your young child is not an especially gifted ballet dancer if she is not an especially gifted ballet dancer. But you should probably not share that belief with her. Don't berate her 3rd grade dance performance. Despite the belief, you should remain supportive, encouraging, and loving toward her in every way.) Even if the awful art case is a counterexample, and I will argue that it is not a counterexample to a particular understanding of CP, it is one that shows that the precise letter of the law, rather than spirit of the law in Clifford's thinking, is flawed. As this case is set up, what's best from a moral point of view, is to believe the artist's work is brilliant. There are several ways to respond to this example. I'll discuss two possible responses to the awful art case. But before I do, I want to stress that even in this situation, there is an important evidentialist insight that should be noted. It is precisely because the evidence tells you that it is morally best to believe against your evidence that obligates you to do so. You are not following a whim, the Ouija board, rumors, or taking a wild guess. Evidence is still absolutely critical in determining what is morally right to believe.

A moral evidentialist should treat the awful art case, and other similar examples, as a straightforward case of conflicting moral obligations. On the one hand, I have a moral obligation to have epistemically justified beliefs, thereby believing the artwork is not brilliant. I also have another moral obligation, namely the obligation to save innocent people and many great works of art. These obligations conflict with one another. In this case, the moral obligation to save the innocent people is stronger than my other moral obligation to believe what my evidence supports. Conflicts between moral obligations are commonplace. Occasionally, for example, I have an obligation to attend a late faculty meeting. Every day, I also have an obligation to come home from work and take my dog out for an enjoyable stroll. On days of late faculty meetings, I have conflicting obligations. The fact that an occasional late meeting is more 'important' does not get me off the hook with my obligation to my dog. In fact, I make arrangements to meet my obligation with the help of a friend who will feed and walk my dog. A similar point could, and should, be made in cases where our doxastic obligations

Sharon Ryan

conflict with, and are outweighed by, other moral obligations. To make this point clearer, consider two ways of thinking about CP2:¹¹

(CP2pf) S has a *prima facie* moral obligation to have doxastic attitude D toward proposition *p* at time *t* iff having D fits S's evidence at *t*.

(CP2a) S has an absolute moral obligation to have doxastic attitude D toward proposition *p* at time *t* iff having D fits S's evidence at *t*.

The awful art objection is only a problem for CP2a. It is not a problem for CP2pf. Whether Clifford would endorse CP2a but reject CP2pf is not clear to me. Again, I'm trying to articulate and defend moral evidentialism in its most plausible form. Whether or not the most plausible form is actually Clifford's is not my concern. CP2pf is a defensible form of moral evidentialism that shares the spirit, if not every detail, of Clifford's views on the ethics of belief. It is also worth pointing out that all of the other alleged counterexamples, if they are actually aimed at moral evidentialism, could be addressed in this way as well, and without having to depend on any of the other criticisms I note.

I think CP2pf, a plausible version of moral evidentialism, survives the inflexibility problem. We should think of doxastic obligations as *prima facie* moral obligations. If you want to live well, you should still make a general habit of having epistemically justified beliefs. Believing what your evidence supports normally has morally good consequences and it demonstrates an honest and virtuous character. Thus, it remains true that you have a moral obligation to believe in accordance with your evidence unless you run into one of these highly unusual situations in which you have strong evidence showing that it would be morally wrong to do so. Working out the details of how to weigh out the moral strengths of our various obligations is beyond the scope of this paper. Let me just note that the weighing is a moral weighing, not a practical or 'all things considered' decision.¹²

A second response, and one that I do not endorse, is to deny that I have a moral obligation to believe the artwork is brilliant. It seems that I could not possibly, if I were in such circumstances, get myself to believe the artwork is brilliant. I'd be able to say I believe it is brilliant, but it is difficult to imagine that I could just, by a mere act of the will, actually believe what I know to be false. The moral motivation won't work. I'm going to be stuck with my belief that the

¹¹ I am grateful to an anonymous referee for helpful suggestions on how to make this point clearer.

¹² I'm thinking here of a view along the lines of W. D. Ross's ethical theory (W.D. Ross, *The Right and the Good* (Oxford: Oxford University Press, 1930).

artwork is awful. If that's psychologically true, and if 'ought' implies 'can,' then it is not true that I have a moral obligation to believe the artwork is brilliant. This alternative protects both CP2pf and CP2a. Unfortunately, it also opens a can of worms for Clifford and for anyone else who thinks we have doxastic obligations of any kind. That's because it seems that all of the doxastic attitudes we have, at any given moment, are not under our voluntary control. As I mentioned, I favor the solution that acknowledges that we can sometimes have, and often have, conflicting moral obligations.¹³ That is, I believe moral evidentialism should be understood as a view that acknowledges that our doxastic obligations are *prima facie* moral obligations. Thus, I think moral evidentialism survives the Inflexibility Problem without having to go into this thorny territory. Nevertheless, we will head there shortly as we take up a second line of objection to CP. However, I want to stress that the Inflexibility Problem can be satisfactorily addressed without going there.

Before moving on to the second line of objection against CP, I hope to have at least shown that the Inflexibility Problem is not as clean and simple as it might initially seem. I believe I have shown more. I believe I have shown that the Inflexibility Problem is solved and that moral evidentialism is a serious, interesting, and strong position in the ethics of belief.

Objection #2: The Involuntarism Problem

Many philosophers reject the idea that we have doxastic obligations, not because they reject evidentialism, but because they believe doxastic attitudes are involuntary responses, not actions that are under our direct, voluntary control. Holding a person responsible for her beliefs, according to these philosophers, makes about as much sense as holding someone responsible for a twitchy eye, the natural color of their hair, or their blood type. We are reminded that 'ought' implies 'can,' and since we cannot control our beliefs, we cannot have any obligations to believe (or disbelieve or suspend judgment.)

The following argument, presented by William Alston¹⁴ and endorsed by many others, is taken by many to be a decisive refutation of the claim that we have obligations, of any sort, about the doxastic attitudes we form. It has also been

¹³ But I do not think we often have moral obligations that conflict with the doxastic obligation to believe what our evidence supports. I think such conflicts are extremely rare. I believe our doxastic obligation to believe what our evidence supports remains undefeated in most circumstances.

¹⁴ William Alston, "Concepts of Epistemic Justification," in his *Epistemic Justification: Essays in the Theory of Knowledge* (Ithaca: Cornell University Press, 1989), 81-114.

Sharon Ryan

used to argue against deontological conceptions of epistemic justification and internalist conceptions of justification, but we will limit our focus to the question of whether Alston's argument undermines the claim that we have doxastic obligations.

The Involuntarism Argument

- (1) If we have doxastic obligations, then we have voluntary control over our doxastic attitudes.
- (2) We don't have voluntary control over our doxastic attitudes.
-
- (3) We do not have doxastic obligations.

Again, I will continue to focus on moral obligations as opposed to purely epistemic, prudential, or other obligations. The Involuntarism Argument, if successful, shows that sentences of the following form are not true:

S ought to believe (or disbelieve or suspend judgment on) *p*.

If the Involuntarism Argument is sound, then it is false that Clifford's shipowner violated a moral obligation when he believed his ship was seaworthy. It is also false that you violate any moral obligation if you refuse to believe ISIS uses brutal tactics after watching a video of them putting a captured Jordanian pilot in a cage and burning him alive. And it is false that you violate any moral obligation if you believe that your neighbor is an awful human being when you have not a shred of evidence to support that belief. If Alston's argument is correct, we have no moral obligations when it comes to forming beliefs.

Alston's argument seems problematic because it does seem true that there are certain claims that we ought to believe, some we should suspend judgment on, and some that we ought to disbelieve. What we believe is central to who we are and what we do. Beliefs are, as Pamela Hieronymi puts it, "a central example of the sort of thing for which we are most fundamentally responsible."¹⁵ We ought, if we are wise, to have a healthy dose of epistemic humility. We ought not be epistemically arrogant. Many people think forgiveness, at least sometimes, ought to be given. But if forgiveness involves beliefs and emotions, and if we regard beliefs and emotions as involuntary in a sense that excludes obligation, it seems that we can never have an obligation to forgive (or not forgive). Moreover, consider the wrongs of racism, sexism, homophobia, anti-Semitism, etc. An important part of why racist behavior is so bad is because we think the beliefs that

¹⁵ Pamela Hieronymi, "Responsibility for Believing," *Synthese* 161 (2008): 357.

lead to the behavior are reprehensible. If we did not think racists could be held responsible for what they believe, then I think we would be much less harsh on racists. If a racist never acts on his or her beliefs, we still think something's very wrong. It is important to acknowledge that Alston's conclusion reaches far beyond a narrow debate among a group of professional epistemologists. Whether doxastic attitudes can be the objects of obligation is enormously important.

Doxastic voluntarists argue against premise 2 of Alston's argument, attempting to show that we do have control over whether we believe, disbelieve, or suspend judgment on *p*.¹⁶ Appealing to a compatibilist view about belief formation, I, myself, argued against premise 2 in an earlier paper.¹⁷ I am no longer sure what I think about premise 2 of Alston's argument. For the purposes of this paper, I am willing to accept that there is a sense in which we do not have direct voluntary control over our doxastic attitudes. What I will argue for here is that the sense in which it seems that we cannot control our beliefs is irrelevant to the 'ought' implies 'can' principle, and that this lack of control is consistent with having doxastic obligations. In the remainder of this paper, I will argue against premise 1 of the Involuntarism Argument.

Premise 1 is apparently supported by the famous, and allegedly obviously true, 'ought' implies 'can' principle. I don't find the 'ought' implies 'can' principle to be obvious. In "Doxastic Compatibilism and the Ethics of Belief," I argued that it (described in four different versions) is false.¹⁸ At this point in my thinking about these issues, I am convinced that the principle is thoroughly ambiguous. Moreover, on its most charitable readings, it is irrelevant to questions about doxastic obligations and therefore does not provide a rationale for Alston's first premise.

Let me begin by noting what I find compelling about one idea that might be what people have in mind when they cite the 'ought' implies 'can' principle. There are some situations where it makes sense to think that doing A is not obligatory because there is a lack of agency or something is impossible for an agent to do. When I think of such cases, I imagine that perhaps we are making use of something that might count as the 'ought' implies 'can' principle. Here is one such

¹⁶ Matthias Steup, for example, has defended Doxastic Voluntarism in numerous articles including "Doxastic Voluntarism and Epistemic Deontology," *Acta Analytica* 15, 1 (2000): 25-56, "Doxastic Freedom," *Synthese* 161, 3 (2008): 375-392, "Belief Voluntariness, and Intentionality," *Dialectica* 65, (2011): 537-599, and "Belief Control and Intentionality," *Synthese* 188, 2 (2012): 145-163.

¹⁷ Sharon Ryan, "Doxastic Compatibilism and the Ethics of Belief," *Philosophical Studies* 114 (2003): 47-73.

¹⁸ Ryan, "Doxastic Compatibilism."

Sharon Ryan

case. Imagine that your friend is seriously injured when a heavy object crashes onto her windshield as she is driving down the road. If someone threw a garbage can off a bridge and it hit her car, it makes sense to hold the hurler responsible. But what if the cause of the broken windshield was a boulder that came loose from a mountainside after a lot of rain? In that case nobody is responsible. No obligation has been violated. Boulders don't have obligations because they are not moral agents. They don't have the ability to plan, respond, or engage in any kind of conscious behavior. It just doesn't make sense to think that the boulder did something morally wrong or failed an obligation. If this is the point of the 'ought' implies 'can' principle, then the principle seems true to me.

But what, exactly, is the principle? And how does it bear on questions about doxastic obligations? Here is one attempt to state the idea:

A: Only moral agents have moral obligations.

Since boulders do not qualify as moral agents, boulders do not have moral obligations and they cannot be held responsible for the effects they have on the world. If A is true, boulders are not morally responsible for breaking windshields. But this doesn't help us to resolve our questions about whether we have moral obligations about what we believe. A does not show that I am not responsible for the beliefs I form. Unlike the boulder, I am a moral agent and A does not get me off the hook for anything I do, including what I believe. So this one very reasonable formulation of 'ought' implies 'can' is totally irrelevant to questions about doxastic obligations.

Let's try another formulation. Perhaps the idea is that in order to be held responsible for an action, it must be at least logically or physically possible for the agent to perform the action. Let's call this B:

B: If S is morally responsible for doing A, then it must be possible for S to do A.

Suppose, for some reason, being born with blue eyes is undesirable. There is nothing I can do about the fact that I was born with blue eyes. So, holding me responsible for the natural color of my eyes is ridiculous. The natural color of my eyes is caused by factors I cannot control. Thus, any claim such as "Sharon morally ought to have been born with non-blue eyes" is ridiculously false. Or suppose I am hanging out in a park with my dog with no particular purpose other than to take a walk and enjoy the fresh air. Imagine that there is a fund-raiser for a great humanitarian cause going on and I get in line to participate. When I reach the head of the line, the organizers inform me that participants must jump over the moon in order to contribute to the cause. I walk away disappointed. I can't have any obligation to jump over the moon. I just plain can't do that. I never have been

able to do so, and I never will be able to do so. But, how does this help show that we don't have doxastic obligations? If the requirements for the fund-raiser were changed so that I had to believe $2+2=4$, I could and I would. This formulation of the principle does not show that we don't have doxastic obligations. Human beings can and do have beliefs – plenty of them.

Going back to the example in the introduction of this paper, I claimed that human babies should not be held morally responsible for their actions. In order to be responsible, a person must be capable of understanding the moral impact of their actions. Human babies lack such comprehension. Thus, they are not responsible even if they do something with terrible consequences. (Rolling over on a pet hamster and injuring it, for example.) If this idea is an 'ought' implies 'can' principle, then I accept it. Perhaps it is:

C: If S is morally responsible for his or her actions, S must understand the moral significance of his or her actions.

But, just like B, C is irrelevant to questions about doxastic obligations, as long as we restrict those obligations to believers who are competent enough to understand the moral significance of their behavior. A, B, and C seem like true principles to me. If that is what people mean when they chant "'Ought' implies 'can!'" then I can appreciate why they find it so compelling. But A, B, and C have no application to questions about doxastic obligations for mentally competent adult human beings.

There are other principles that seem relevant to Alston's argument, and they are discussed in the ethics of belief literature, but they are much less plausible. And, they take us away from any ideas that obviously connect to 'ought' implies 'can.' So far away, in fact, that I will consider them without even trying to make the case that they are 'ought' implies 'can' principles. Here's one idea one might appeal to in backing up premise (1):

D: If S is responsible for doing A, then S must do A intentionally.

D requires not only that the responsible agent be able to do A, but that she do A, and do so intentionally. One might use something such as D to argue against doxastic obligations. Although human beings can and do believe things, there is some reason to think that the formation of a belief is not intentional.¹⁹ When I pay attention to what I am doing and notice that I am typing a sentence, the belief

¹⁹ Some philosophers argue that belief formation is intentional or deliberate. Although I will not take up any of those arguments in this paper, I wish to acknowledge that there are such arguments and they are worthy of serious consideration. I am assuming, just for the sake of the argument, that belief formation is not intentional, voluntary, or deliberate.

Sharon Ryan

that I am typing a sentence just shows up. I don't cause myself to believe $2+2=4$ and I don't cause myself to believe Barack Obama is President. When my cat, Diego, awakens me by knocking things off my dresser, the belief that "Diego is up to no good again" is forced upon me by my perceptual experience. I don't intentionally decide to form the belief and then believe. My beliefs are formed without any intention of mine playing the right kind of causal role. So far, so good, but D is false. We hold people responsible for unintentional actions all the time. A driver who is distracted by texting and unintentionally runs over a cyclist is morally responsible for running over the cyclist. Suppose the driver didn't even see the cyclist and had no idea what happened when she felt the bump and heard the thud. Nevertheless, she ought not to have hit the cyclist. And, she's morally responsible for doing so.

One might try to salvage D, or something similar to D, by pointing to the many actions the driver did do intentionally that help explain why we hold her responsible for running over the cyclist. Perhaps she intentionally decided to start texting. That decision led to the unfortunate consequence. Thus, we can hold her responsible for running over the cyclist because she was texting intentionally. But we could say the same thing about belief formation (assuming for the sake of the argument that belief formation is not intentional.) When I wind up with a doxastic attitude, it is not caused by magic. I never find myself with a belief and think, "Wow, where did that come from?" My coming to believe x (according to those who claim that forming beliefs is involuntary) is caused by many psychological processes such as deciding to think about particular questions, reading, concentrating, observing, weighing my evidence, listening to the world around me, checking my sources, and so forth. These actions are typically intentional. So, even if we accept D, and discount the texter as a serious problem for D, we can make the same claims about doxastic attitudes that we make about the texter. Dustin Olson has developed a view of epistemic agency along these lines. While rejecting doxastic voluntarism, Olson argues that it is our ability to develop and refine our belief-forming methods and practices that provides us with a way of making sense of epistemic responsibility. According to Olson,

We can be held responsible for our beliefs because there are things we can do that can affect them – *a fortiori* we do have the right kind of control to allow for epistemic duties.²⁰

²⁰ Dustin Olson, "A Case for Epistemic Agency," *Logos & Episteme. An International Journal of Epistemology* VI, 4 (2015): 449-474.

If D is what is at work in the Involuntarism Argument, and I think that it or something very much like it is, the argument fails. D seems false to me, and the only way to salvage it opens a door that salvages doxastic attitudes that are not formed intentionally.

But perhaps I am still missing the point. Perhaps the reason we want to hold the texter responsible for hitting the cyclist is because although she hit the cyclist unintentionally, hitting a cyclist (or driving carefully or recklessly) is the sort of action that can be done intentionally. So, maybe the idea is that in order for A to be open for moral evaluation, A has to be the sort of action that can ever be an intentional action. Driving carefully is such an action, but the argument alleges, forming a belief is not. Forming a belief is never, we can suppose, an intentional action. Thus, consider:

E: If S has a moral obligation to do A, then A must be the kind of action that can be done intentionally.

Assuming, for the sake of the argument, that we can never form beliefs intentionally, E would generate the result Alston wants. Although E is relevant, it is false. There are a lot of things for which we can be held responsible that we can't do, or control, intentionally. We are responsible for having healthy cholesterol levels (if we want to be healthy and stay healthy enough to take care of children we are responsible for bringing into the world). We don't intentionally and directly control our cholesterol numbers, but we are responsible for keeping them within a good range. If we don't, there can be horrible moral consequences. Of course, we can do something to affect those numbers. We can decide to eat certain foods and avoid eating others, we can exercise, and if all else fails, we can take medication. But, it is the causal effects of those decisions that control our cholesterol levels. Intentional acts of will can't do the work. Try as you may, merely wanting or deciding to lower the numbers will be ineffective. But, again, a similar point can be made about our beliefs.²¹ They can have serious moral consequences, and there is a lot that we can do in our epistemic practice that will have an effect on what and how we believe. Again, the ability to control, at will or intentionally, is not necessary for fair moral attributions.

Perhaps I have been trying too hard to come up with a rationale for premise 1. Maybe premise 1 does not have much of a rationale. Perhaps Alston and other advocates think it does not need any further support because is just plain obvious. Perhaps defenders of the Involuntarism Argument think it is self-evidently true that we must have immediate, voluntary control over any action that is subject to

²¹ See Dustin Olson's "A Case for Epistemic Agency" for a developed defense of this idea.

Sharon Ryan

moral appraisal. There is a certain kind of control that we do seem to lack over our doxastic attitudes and I am willing to call that voluntary control. I can't form a belief just because I feel like it. Interestingly this isn't always the case with ordinary actions. For example, right now, if I want to get up from my desk and walk around and think for a while, I can just do it. In contrast, if I want to believe that I am skiing in Norway right now (as I am fully aware that I am writing a philosophy paper in Morgantown, West Virginia), I can't just do it. I can *imagine* skiing in Norway right now, but I can't actually believe it. I just don't have that kind of control over my beliefs. I believe George Washington was the first President of the United States. If a friend asks me to disbelieve that claim, I can't. If I could save a million lives by believing that Thomas Jefferson was the first President of the United States, I'm going to be letting a million or more people down. I can't just do it. I'm, as Richard Feldman puts it, 'at the mercy of my evidence.' But how is that supposed to show that we do not have doxastic obligations?

We are responsible for other actions that are not under this sort of immediate voluntary control. Imagine that Johnny's school play starts in 10 minutes and Johnny's dad is coming. Johnny's dad has a lot of responsibilities. One of them is to be sober at the play. But Johnny's dad is drunk as a skunk. He can't now be sober at the play. He can't right now decide to be sober at the play and execute that decision for a million dollar pay out. He can't do it if he realizes it would be best for his relationship with his son. He might want to be sober, but he can't, just by willing it, be sober for the play. He doesn't have that kind of control. Nevertheless, he still ought to be sober at his son's play. Again, we can trace his obligation to be sober at the play back to other things he could and did control. He decided to drink and just started drinking. But, again, the same is true of beliefs. I can decide to pay attention to all of my evidence, I can decide to take counterevidence into account, I can decide to take courses that improve my reasoning skills, I can decide to buy books and read them, etc. And, with respect to Johnny's dad, we look forward to the consequences, to see that they are serious and morally important. The same can be said of beliefs. Remind yourself of Clifford's shipowner or the effects of racist beliefs.

Even if we lack control at the final step in belief formation, there is a lot leading up to the final step, and following that final step, that we do control. If we can hold people responsible for actions under such circumstances, why can't we hold people responsible for their beliefs?

Suppose I decide to jump off a diving board into a pool. Once I am up in the air, I realize that I am wearing my friend's \$1,000 (non-water proof) watch that

was a gift from her now deceased mother. I don't want to ruin her watch. I didn't mean to ruin her watch. At this point, I can't make an effective decision to keep the watch safe - not for all the money in the world, not for the sake of the friendship, not for anything. I'm at the mercy of gravity, and I'm responsible for ruining my friend's watch. Why am I responsible? I should have taken a second to reflect on what I was wearing before I jumped. I was careless. Plus, I am ruining, for no good reason, a prize possession of my friend's. The causes were under my control and the consequences are morally bad.

My point is that we do hold people responsible for doing things they do not have immediate voluntary control over. In many ordinary situations, we hold people responsible for being sober at their kid's play, having healthy cholesterol levels, not ruining their friend's watch, even when those achievements cannot be obtained by a mere act of the will. Even if beliefs are not under our direct control, even if the particular doxastic attitude formed is involuntary or unintentional, there are excellent reasons to hold us responsible that are perfectly analogous to our responsibility for many morally significant actions. Furthermore, because of the enormous moral importance the effects of our beliefs can have, and the enormous effect our epistemic practices have on our beliefs, it seems appropriate to hold us responsible for what we believe, even if, at the exact moment of belief formation, we are at the mercy of our interpretation of the information we have in front of us.

I believe the analogy between beliefs and morally obligatory, yet involuntary and unintentional, actions is strong. Thus, I think we should treat them similarly. If we are willing to grant that Johnny's dad ought to be sober at the play and that the texter ought not have hit the cyclist, we ought to accept that I ought to believe ISIS is brutal and Clifford's shipbuilder should not have believed his ship was seaworthy. I accept all of the above. However, there is one more convincing reason to reject the Involuntarism Argument and it does not depend on treating beliefs and actions analogously. Philippe Chuard and Nicholas Southwood challenge premise 1 by noting that we make normative judgments about other attitudes that are as involuntary as are beliefs.²² For example, under the right set of circumstances, it makes sense to say, "S ought not be angry at me." The circumstances might be that I did not do what S thinks I have done and I have proven to S that I did not do it. After seeing the situation in this new light, S ought not be angry. And the 'ought' is a straightforward moral ought. Why not say

²² Philippe Chuard and Nicholas Southwood, "Epistemic Norms without Voluntary Control," *Nous* 3, 4 (2009): 599-632.

Sharon Ryan

similar things about beliefs? If we do, we have another good reason to deny premise 1.

The Involuntarism Argument is weak. Even granting, for the sake of the argument, that beliefs are involuntary, this argument does not show that we do not have doxastic obligations.

Doxastic Obligations, Moral Responsibility, and Wisdom

At this point, I believe I have successfully defended a version of moral evidentialism from two serious and widely accepted arguments. In closing, I'd like to suggest that we need doxastic obligations both to provide an adequate explanation for the degree to which we hold people morally responsible for reprehensible behavior and to understand what it takes to be wise.

Consider, for example, the moral judgments we make about people who are willing to blow up buildings with lots of innocent people inside. Consider how we feel about practicing racists, sexists, and other haters. It is not just their actions that are so disturbing, but the crazy ideas behind those actions. If we do not hold people responsible for what they believe, and for and how they arrive at their beliefs, then I think we are too harsh in our moral judgments of behavior that results from those beliefs. If we are unable to hold a sexist person responsible for his or her sexist beliefs, then it is difficult to hold such people responsible, to any significant degree, for their sexist behavior. It seems harsh to blame a person for his or her actions if she is not responsible for the beliefs that lead to those actions. After all, there is some virtue in acting in consistency with your beliefs. When we consider actual cases of psychologically normal people doing morally bad things, most of them suffer from having unjustified beliefs. Most psychologically normal people don't set out to do something they regard as morally wrong. Most psychologically normal people act on the basis of what they believe to be right. It is there, when people are thinking about (or not thinking about) what they ought to do, that a lot of effort should be demanded.

Doxastic obligations are also an important aspect of wisdom. Wise people ought to, among other things, believe in accordance with their evidence. Wise people ought to have appropriate emotional responses. Beliefs and our emotions are not under our immediate voluntary control, and yet they are, perhaps more than anything else we do as human beings, of enormous moral importance. If what I have argued for in this paper is correct, the most compelling arguments

against doxastic obligations fail and we have good reason to believe that we do have doxastic obligations.²³

²³ I would like to thank two anonymous referees and the participants at the November 2014 Cornell Workshop on the Ethics of Belief for helpful comments on earlier drafts of this paper.

MORAL REASONS FOR MORAL BELIEFS: A PUZZLE FOR MORAL TESTIMONY PESSIMISM

Andrew REISNER and Joseph VAN WEELDEN

ABSTRACT: According to moral testimony pessimists, the testimony of moral experts does not provide non-experts with normative reasons for belief. Moral testimony optimists hold that it does. We first aim to show that moral testimony optimism is, to the extent such things may be shown, the more natural view about moral testimony. Speaking roughly, the supposed discontinuity between the norms of moral beliefs and the norms of non-moral beliefs, on careful reflection, lacks the intuitive advantage that it is sometimes supposed to have. Our second aim is to highlight the difference in the nature of the pragmatic reasons for belief that support moral testimony optimism and moral testimony pessimism, setting out more clearly the nature and magnitude of the challenge for the pessimist.

KEYWORDS: evidentialism, moral epistemology, moral testimony,
reasons for belief

Introduction

Testimony is commonly thought to provide normative reasons for belief, at least in favourable circumstances. A standard favourable circumstance is one in which an individual has sufficient reason to believe that some other agent is reliably better positioned than she is to have or arrive at true beliefs about a particular domain of inquiry.¹ Call such a person an *expert* about the relevant domain. The domain in question may be empirical, such as quantum physics or plumbing. The domain may also be theoretical, such as mathematics or logic. In each case, expert testimony in support of particular claims in those domains provides normative reasons for non-experts to believe those claims.² Perhaps it would be epistemically better in some way for a non-expert to study up on plumbing or number theory, in order that she may arrive at all the important truths in the vicinity through her own powers of reasoning. Still, due to the cognitive limitations of humans, in

¹ And assume that this agent is also more likely to be right than I am for each individual belief on which we disagree.

² We shall set aside the question of whether, and how, other experts ought to react to expert testimony.

general, it is not the case that non-experts ought to study up on each particular domain of inquiry. In many cases non-experts not only have a normative reason to defer to expert testimony, but also ought all-things-considered to do so. Some cognitive division of labour is both desirable and inevitable.

Given this, it is perhaps surprising that various philosophers³ have argued that there is something wrong with forming at least one class of beliefs – the class of moral beliefs – on the basis of expert testimony. Some of these same philosophers are prepared to grant that there are truths about the moral domain, and also that there are identifiable moral experts. The strongest anti-deference positions are versions of what we will call *moral testimony pessimism*, according to which expert moral testimony does not provide us with normative reasons for belief. There can be weaker versions, which both admit that expert moral testimony is reason-providing and insist that deference to moral testimony is nevertheless something against which there is a standing *pro tanto* reason. Roger Crisp, in particular, sometimes seems to be arguing for the weaker position, although he may ultimately support moral testimony pessimism as characterized above.⁴ We are agnostic about the weaker claim for the purposes of this paper. Our argumentative focus is moral testimony pessimism understood as the stronger claim.

The combination of the views that there are moral truths, that there are identifiable moral experts, and that expert moral testimony is not reason-providing, is what most interests us in this paper. This is because we think that this combination of views is difficult to sustain, and also because the difficulties here connect with important issues in the broader literature on normative reasons for belief. In the broader literature, there is a widespread assumption that *strict normative evidentialism*, or some closely related view, must be true. Strict normative evidentialism (‘evidentialism’ for the rest of the paper) is the view that

³ This includes such recent figures as Alison Hills (“Moral Testimony and Moral Epistemology,” *Ethics* 120, 1 (2009): 94-127), Robert J. Howell (“Google Morals, Virtue, and the Asymmetry of Deference,” *Nous* 48, 3 (2014): 389-415), Robert Hopkins (“What is Wrong With Moral Testimony,” *Philosophy and Phenomenological Research* 74, 3 (2007): 611-634), Sarah McGrath (“The Puzzle of Pure Moral Deference,” *Philosophical Perspectives* 23, 1 (2009): 321-344), and Philip Nickel (“Moral Testimony and its Authority,” *Ethical Theory and Moral Practice* 4, 3 (2001): 253-66).

⁴ Roger Crisp, “Moral Testimony Pessimism: A Defense,” *Proceedings of the Aristotelian Society Supplementary Volume* LXXXVIII (2014): 129-43.

all normative reasons for belief are or are constituted by evidence for the contents of the beliefs for which they are reasons.⁵

While we are convinced that the strict version of evidentialism has been shown to be untenable on a variety of fronts, we feel it is important to observe that one's commitments with respect to evidentialism have ramifications for one's view about whether there can be reasons for having or forming moral beliefs based on expert testimony.⁶ There are a variety of ways in which evidentialism might be used to defend moral testimony optimism, the view that moral expert testimony does provide normative reasons for moral beliefs, so that its rejection might be seen as an important step on the road to a defence of moral testimony pessimism.

Even with our anti-evidentialist sympathies, we will argue that moral testimony pessimists have the tougher row to hoe. We shall argue that standard arguments for pessimism commit their proponents to a particularly controversial kind of pragmatism about reasons for belief. This commitment, we argue, is self-undermining. A liberal form of pragmatism about reasons for belief may provide *pro tanto* reasons against deferring to moral expert testimony, but in many cases it will provide stronger positive reasons for deferring. As noted above, some authors who identify themselves as pessimists are committed only to the weak thesis that there are some standing normative reasons for not deferring to moral expert testimony, but allow that these reasons may be outweighed in many or even most cases by the reasons for deferring to moral expert testimony. We call this version

⁵ Closely related views allow some flexibility for logical truths and certain kinds of reasons that count *a priori* towards the truth of a belief, when evidence does not seem to be quite the right kind of way to describe the truth-indicating relation. We shall not distinguish between these closely related views and strict normative evidentialism proper in this paper, although the differences may be important in other contexts.

⁶ For cases made directly against strict normative evidentialism, see: Miriam McCormick, *Believing against the Evidence: Agency and the Ethics of Belief* (New York: Routledge, 2015), Andrew Reisner, "Evidentialism and the Numbers Game," *Theoria* 73, 4 (2007): 304-316, Andrew Reisner, "The Possibility of Pragmatic Reasons For Belief and the Wrong Kind of Reasons Problem," *Philosophical Studies* 145, 2 (2009), 257-272, Andrew Reisner, "Leaps of Knowledge," in *The Aim of Belief*, ed. Timothy Chan (Oxford: Oxford University Press, 2013), 167-183, Andrew Reisner, "A Short Refutation of Normative Evidentialism," *Inquiry: An Interdisciplinary Journal of Philosophy* 58, 5 (2015): 477-485, doi: 10.1080/0020174X.2014.932303, Asbjorn Steglich-Petersen, "Does Doxastic Transparency Support Evidentialism?" *Dialectica* 62, 4 (2006): 541-547, Asbjorn Steglich-Petersen, "Transparency, Doxastic Norms, and the Aim of Belief," *Teorema* 32, 3 (2013): 59-74, and Sarah Stroud, "Epistemic Partiality in Friendship," *Ethics* 116, 3 (2006): 498-524.

of pessimism ‘*pro tanto*-ism.’ It is our contention that while *pro tanto*-ism may be plausible, it is in a certain sense uninteresting.

Although evidentialism only supports optimism, both optimism and pessimism may be supported by non-evidential moral reasons for belief. In light of this, we shall emphasise that there is a distinction between two types of moral reasons for belief – those that by their nature are consistent with the alethically generated reasons and those that are not – that may also be found more generally amongst pragmatic reasons for belief. Indeed, it is our claim that pessimists are, perhaps unwittingly, committed to the existence of pragmatic (specifically moral) reasons for belief that are contrary to the evidence. We have no objection to the existence of these reasons, but we feel that this point has been underappreciated by pessimists, who do not typically present their position as involving such a strong commitment about the nature of reasons for belief themselves. This last point makes moral testimony an interesting test case for thinking about pragmatic reasons for belief more generally.

Our aims in this paper are twofold. We first aim to show that testimony optimism is, to the extent such things may be shown, the more natural view about moral testimony. Speaking roughly, the supposed discontinuity between the norms of moral beliefs and the norms of non-moral beliefs, on careful reflection, lacks the intuitive advantage that it is sometimes supposed to have. The second aim is to highlight the difference in the nature of the pragmatic reasons for belief that support testimony optimism and testimony pessimism, setting out more clearly the nature and magnitude of the challenge for testimony pessimists.

1. Gillian’s Island

We begin by presenting a case that brings out the puzzling nature of the moral testimony pessimist’s claim.

Gillian has set sail for what she intends to be a three hour tour. Along the way she encounters unexpectedly rough weather, and is soon blown off course. Her ship runs aground on an uncharted desert island.

After disembarkation, Gillian learns that the inhabitants are at war both with each other and with all who visit the island. Encountering one of the inhabitants, she believes that her life is now in danger. She hears more inhabitants coming. She calculates that she can run away with little risk to her life, irrespective of the inhabitant’s intentions, but only if she shoots the inhabitant with the ship’s harpoon gun. It is also possible that this inhabitant means her no harm. How does she decide what to do?

Gillian is by nature a morally concerned person, and she wants to do not just what is best for herself, but what is morally right. However, she has no idea what that is. Nothing in her past experience, which has for several years involved little more than being the first mate on Hawaiian pleasure cruises, has given her the tools to make this decision. She lacks the time for adequate reflection on her present situation. She also foresees that she will, until she escapes the island, be faced again and again with new and difficult moral choices, on which she will not have time to reflect in a reasonable and truth-conducive way.

Fortunately, Gillian is not entirely unprepared for such a scenario. She recalls having learned in a university philosophy course that moral experts are people of whom the following is true:

[*Moral Expert*] Someone is a moral expert if the probability that each of her moral beliefs is true is significantly higher than the probability of a non-expert's beliefs being true.

Gillian knows that she is a moral non-expert. Fortunately, prior to departure, she brought three special e-readers aboard the ship. Two of the e-readers have a list of the moral beliefs of a particular moral expert,⁷ and one has a list of the moral beliefs of a non-expert. All of the e-readers can scan Gillian's overall belief state in real time and can then flash the contents of the appropriate moral belief from the e-reader onto a HUD on her eyeglasses.⁸ Gillian knows that because the expert's beliefs will be fed to her in a state of great agitation and susceptibility, she will most of the time give significant weight to the expert's beliefs in the formation of her own moral beliefs.

Gillian can only bring one e-reader with her as she explores the island, and thus will have only one with her before disembarkation. It seems wrong to her not to choose one of the first two, but as she weighs the matter, a complication arises. She notices a small brochure attached to the third e-reader. The brochure explains that the beliefs contained therein are those that are entailed by Gillian's most foundational moral views, and moreover are those that she would in fact arrive at given more time to deliberate.

The matter becomes further complicated, as the brochure goes on to explain that the final e-reader is not in fact an e-reader at all. It is just a fancy container for a pill. If consumed, the pill will accelerate Gillian's moral reasoning, when faced with a crisis situation, to a speed at which she can, and will, form the moral

⁷ That is to say that each e-reader contains a list of one moral expert's beliefs, and that there is a different expert's beliefs on each of the two e-readers.

⁸ Alternatively, one could treat the e-readers as having access to the same evidential field as Gillian.

beliefs that she otherwise would have arrived at much more slowly under ideal deliberative circumstances. Gillian must now decide what to do. She can take no e-reader with her, she can choose one of the two expert e-readers, or she can choose to take the third e-reader cum pill case.

2. E-Readers and Pills

In thinking about what Gillian ought to do, it will be helpful to begin with three observations. First, from the point of view of gaining true moral beliefs, Gillian ought to prefer the expert e-readers to the reasoning pill. Between the two expert e-readers, Gillian has no reason to prefer one to the other. This latter point brings out an interesting feature of expert testimony in general. It is in an important sense impersonal. As moral experts are defined in the Gillian's Island example, their testimony about any individual claim is more likely to indicate the truth than any non-expert's. In its non-philosophical usage, the phrase 'moral expert' is often used to denote various kinds of individuals who think carefully about moral questions and are in a position of moral authority, for example priests and hospital bioethicists. Whether, or to which sort of moral expert – in the popular sense of the expression – an individual actually defers is likely to depend on important details both about the individual's circumstances and the expert's. It is an open question whether the categories of people treated as moral experts in the popular domain are moral experts in the sense that is discussed in the philosophical literature. The moral experts of philosophy are not gurus or wise folk, they are just people who are more likely to have true moral beliefs than non-experts and to testify to those beliefs sincerely.

A second and more important observation is that if moral testimony pessimists are right, one ought to have differing views about which e-reader Gillian ought to take, when there is a choice between expert and non-expert e-readers on the one hand and expert e-readers and a moral reasoning accelerator pill on the other hand. Moral testimony pessimists would reject using any e-reader, expert or not, rather than relying on one's own moral reasoning. However, when there is a pill that can rapidly accelerate one's reasoning, with the result that one will form the identical beliefs that one would have formed taking a non-expert e-reader, they would presumably regard it as permissible to take the reasoning accelerator pill, as it involves no deference.

This brings us to the third observation, one that concerns an extension of the original case. We can stipulate that Gillian's individual moral beliefs will each be less likely to be true if she does not take the pill than they would be if she took it. We can assume that this is because her reasoning is rushed and incomplete

when she reasons under pressure with only her normal cognitive resources. Presumably pessimists would want to say in this case that Gillian ought to take the reasoning accelerator pill, as it improves her moral reasoning without necessitating her deferring to the beliefs of others.

Taking the pill is an action, and thus the reason that supports taking the pill is an instance of a reason for doing rather than for believing. We can imagine yet another alternate scenario in which Gillian has a choice between a reasoning accelerator pill and a susceptibility to expert suggestion pill. The latter will cause Gillian to believe all expert testimony without any intermediate reasoning. Without it, Gillian will only rely on her own moral reasoning. Choosing between the two pills is also an action, and the direct reasons that govern it are reasons for action.

The availability of the two different pills puts Gillian in an interesting position. She has the ability to cause herself to have a significantly alethically improved set of moral beliefs, by taking the susceptibility to suggestion pill. She has the ability to cause herself to have a somewhat alethically improved set of moral beliefs by taking the reasoning accelerator pill. Assuming for the sake of argument that having true moral beliefs (or moral beliefs that are more likely to be true) will cause Gillian to act in a morally preferable way over having fewer true (or likely to be true) moral beliefs, there is a straightforward moral reason to take the susceptibility to suggestion pill. This moral reason is, again, a reason for action.

If, as we shall argue later,⁹ the pessimist must claim that there are non-alethic moral reasons for belief, there is an apparent non-alignment between the moral reasons for belief themselves and the moral reasons for action, even where the actions in question are those that will partially determine which beliefs one has and how one forms them.

An alternative picture would be one on which the moral reasons for action – for choosing which pill to take – are wholly parasitic on the reasons for having or acquiring the relevant moral beliefs. For now, we will be assuming that the parasitic hypothesis is correct. When we speak of Gillian's choosing to take a particular e-reader, or pill, we are assuming that the reasons are derivative from her reasons to have or acquire certain moral beliefs. In section 5, we shall discuss why adopting the non-alignment hypothesis rather than the parasitic hypothesis is problematic for the pessimist.

⁹ See sections 4 & 5.

3. Gillian among the Pessimists

From one perspective it seems obvious what Gillian ought to do. According to the case as described, each of the first two e-readers is programmed to produce the moral judgments of some particular moral expert. Although our moral experts need not be infallible and may disagree amongst themselves, a weak condition on being a moral expert is that one is more likely than non-experts to arrive at true moral beliefs, after having engaged in careful moral deliberation. As Gillian is a non-expert, should she choose to take the pill rather than to defer to either of the expert e-readers, she would thereby be deliberately choosing a course of action that increases the likelihood that she will have false moral beliefs.

For the sake of the example, let us assume that acting on true moral beliefs is more likely to result in performing the morally right action than acting on false moral beliefs. Let us also assume for the sake of the example that the aggregate result of acting on true moral beliefs is such that there are no grounds for objecting morally to adopting the policy of acting on individually true moral beliefs more of the time rather than less of the time. If we take these assumptions on board, and we assume that Gillian is enkratic, it is difficult at first blush to see how we could reasonably endorse Gillian's doing anything other than taking one of the expert e-readers.

If there are *pro tanto* reasons for Gillian to take an expert e-reader with her and to acquire the beliefs to which it testifies, it is so in part because the testimony of moral experts is good evidence for the truth of the claims for which it is testimony. Taking an expert e-reader, and forming beliefs on the basis of what it says is the case, is the most successful way for Gillian to conform to the epistemic norm of believing in accordance with the evidence.

However, possessing true moral beliefs is surely not desirable for epistemic reasons alone. Because of the special connection between true moral belief and morally right action, we seem to have special (that is to say, moral) reasons to be concerned with getting the facts about morality correct. Since we should certainly aim at doing what is morally right and avoiding doing what is morally wrong, we should also aim to have true moral beliefs. Given that there thus seem to be powerful moral, as well as epistemic, reasons to choose an expert e-reader, it is puzzling how anyone could recommend that Gillian choose otherwise. Yet various philosophers in the recent literature have defended pessimism and made arguments that seem to commit them to giving such an answer in Gillian's case.

There are several strategies that moral testimony pessimists have taken. In assuming both moral cognitivism and the existence of identifiable moral experts, our case already takes some of these off the table. Sarah McGrath has argued that

it is much easier to account for what is suspect about moral deference on a non-cognitivist picture, according to which there are strictly speaking no moral truths to be an expert about (or to defer to the expert about).¹⁰ Michael Cholbi has argued that there are no identifiable moral experts, so the issue of whether to defer to their testimony does not really arise.¹¹ We will here concern ourselves with arguments against moral deference from those who grant these assumptions. There are two broad strategies left to the opponent of moral deference. The first points to epistemic reasons against deferring to the testimony of moral experts, the second to moral reasons. For either of these general pessimist strategies to gain any traction, it must first be granted that some normative reasons for belief are pragmatic. If there is any interesting problem of moral deference, that is, it is because pragmatism about reasons for belief has some plausibility.

4. Does What Gillian Knows, But Does Not Understand, Hurt Her?

Alison Hills claims that we must distinguish between moral *knowledge* and moral *understanding*.¹² The former can be acquired purely by means of deference to a moral expert's testimony, the latter cannot. Moral understanding purports to be a more demanding notion than moral knowledge, and calls for "a grasp of the relation between a moral proposition and the reasons why it is true."¹³ Hills thinks that the non-expert who arrives at true moral beliefs on account of her deferring to expert testimony does not have this. Because, she claims, understanding is more epistemically valuable than mere knowledge, such an agent is epistemically deficient, even though she is in possession of the same moral knowledge as the expert. As such, there is at least some epistemic reason against deference to moral experts.

There may be something to be said for the view that a grasp of the reasons why some belief is true confers some additional epistemic value on the holding of that belief.¹⁴ Insofar as the e-readers in the case above only provide the moral

¹⁰ McGrath, "The Puzzle of Pure Moral Deference." It is worth noting, moreover, that if moral sentences do not express beliefs but rather non-cognitive attitudes, the pessimist also need not endorse pragmatism about reasons for belief.

¹¹ Michael Cholbi, "Moral Expertise and the Credentials Problem," *Ethical Theory and Moral Practice* 10, 4 (2007): 323-334.

¹² Hills, "Moral Testimony and Moral Epistemology."

¹³ Hills, "Moral Testimony and Moral Epistemology," 101.

¹⁴ Although whether there is anything to such a *grasp* of the relevant reasons beyond knowledge that they are the reasons is itself debatable. For an argument that to possess 'moral understanding' is just to have some additional moral knowledge see Amber Riaz, "Moral Understanding and Knowledge," *Philosophical Studies* 172, 1 (2015): 113-128. An interesting

expert's judgments, without the reasons that ground them, perhaps something important will be missing from the deferential agent's final epistemic position.

However, so far this is only to say that true moral beliefs arrived at by deference are not as epistemically valuable as true moral beliefs arrived at in a way that grants understanding. It may well be epistemically preferable to have moral understanding and not just moral knowledge. Not even the optimist about moral testimony need deny this. Indeed, one might say the same about quantum physics or plumbing. There seems to be something epistemically admirable about retaining a grasp of both the relevant facts (about morality, or physics, or plumbing) and of what explains them.

Suppose it is granted that the epistemic value of understanding is greater than that of mere knowledge. In order to justify the claim that it is ever all-things-considered better not to defer to moral experts, when they are available and identifiable, one would need a more general ranking of the comparative epistemic value of different states (true belief without knowledge, mere knowledge, understanding, false belief, etc.). For the pessimist's case to be successful, she must show that small gains in moral understanding, in combination with more significant losses in true moral beliefs, dominate more significant gains in moral knowledge in combination with equally significant gains in true moral beliefs.

In the case of Gillian, if she does not opt to defer to an expert by taking one of the e-readers, she is knowingly making it less likely that the moral beliefs she will arrive at will be true: by not taking the expert e-reader, Gillian will presumably fail to acquire many of the beliefs that evidence requires her to have. Insofar as the requirements of evidence are or provide normative reasons, Gillian will not have many of the moral beliefs that she has strong reason to have.

Philosophers who wish to defend the view that moral understanding is more epistemically valuable than mere moral knowledge may be committed to one of two claims, or to both. The first claim is that when both are attainable, there is more epistemic reason to seek after understanding than after mere knowledge. The second claim is that it is more epistemically valuable (without any explicit commitment to there being special reasons) to have understanding than mere knowledge.

It is difficult to see how either of those claims, or a combination of them, can be leveraged into an argument for an interesting version of pessimism. Moral understanding, which requires that one grasps the reasons why a belief is true,

general discussion of these issues appears in Daniel Star's *Knowing Better: Virtue, Deliberation, and Normative Ethics* (Oxford: Oxford University Press, 2015).

may be in the relevant sense a higher epistemic good than mere moral knowledge. But, a prerequisite for moral understanding is that one have true moral beliefs.¹⁵

This makes Hills style pessimism epistemically odd. One cannot have moral understanding without true moral beliefs, or perhaps mere moral knowledge.¹⁶ Yet, one is required not to form moral beliefs in accordance with the best indicators of moral truth. In fact, one is required knowingly to adopt moral belief formation procedures that are less likely to yield true beliefs than those that incorporate moral testimony. In effect, Hills is suggesting a set of epistemic norms that at once values understanding while undermining a necessary condition – the having of true beliefs – for possessing it.

The Hills style case against moral deference becomes even harder to make when one adds in the moral reasons that appear to weigh in favour of deferring to the testimony of moral experts. Here is what looks like a plausible moral principle: It is morally wrong to knowingly make oneself more likely to commit moral wrongs.¹⁷ Suppose, like Gillian, you are a moral non-expert. You know that you are more likely to do wrong than an expert would be. Although you do not know exactly when or how you will make mistakes, you do know that, if you follow your own lights, you are very likely to go wrong eventually in some cases in which the moral expert would not err. This is just what it is to be a non-expert. Although you cannot in advance identify the particular wrong actions, you can be confident that in opting never to defer you are thereby increasing the probability that you will perform certain morally wrong acts. If you neglect to defer to the testimony of moral experts, you thereby seem to be proceeding in a way that is itself morally wrong.

We need not deny that acting rightly on the basis of moral understanding confers greater moral value on an action than it would have, were it performed

¹⁵ This point is underscored by Daniel Star's explanation of how there can be both moral experts and non-experts that have moral knowledge. According to Star, the difference is that moral experts know the genuine explanations of moral truths, whereas non-experts with moral knowledge are sensitive to moral evidence, even if they are not always aware of the explanations. See his *Knowing Better*.

¹⁶ Just understood as being moral knowledge absent understanding.

¹⁷ Both Aristotle and Hastings Rashdall are committed to a claim of this sort, although they do not put it probabilistically. See Aristotle, *Nicomachean Ethics*, trans. Roger Crisp (Cambridge: Cambridge University Press, 2000), 1111b. See also Hastings Rashdall, *The Theory of Good and Evil*, vol. 1 (London: Oxford University Press, 1907), 76, Hastings Rashdall, *Is Conscience an Emotion?* (Boston: Houghton Mifflin, 1914), 33, and Hastings Rashdall, *Ethics* (London: T.C. & E.C. Jack, 1913), 69-70. David Enoch has recently made use of much the same principle, this time in explicitly probabilistic terms, also while defending moral deference against its critics. See his "A Defense of Moral Deference," *The Journal of Philosophy* CXI, 5 (2014): 241-3.

without moral understanding. We can even concede that such considerations may provide us with *pro tanto* reasons against deferring to moral experts. However, it is excessively high-minded to suggest that foreseeably setting oneself up to perform wrong actions on a more frequent basis than necessary, in order that one might from time to time perform right actions on the basis of moral understanding, is itself morally commendable. It is a deeply unsettling feature of Hills's view that it seems to treat doing the right thing as being, in the scheme of things, rather less important than one's reasons for having done it. As it is typically others who are harmed by one's wrongdoing, this amounts to a morally problematic fetishisation of how an agent comes to acquire their moral knowledge (or justified true moral beliefs), at the cost of foreseeable harm to individuals other than the agent herself.

5. A Morally Worthy Argument?

Others in the literature argue, more promisingly, that there are distinctive moral reasons not to defer to moral experts.¹⁸ On this kind of view,

the norm that excludes adopting moral testimony is itself rooted in moral considerations i.e. considerations of the kind that ground first-order moral claims.¹⁹

If this were the case, we would be left with moral reasons on both sides, and the issue would become one of weighing these against each other. We shall proceed to discuss just what these competing moral reasons would have to look like in relation to the evidential reasons.

Amongst those who find something morally suspect about moral deference, there are two broad accounts of what that something is. The first is motivated by similar considerations to those to which Hills was responding. According to one influential view, morally worthy actions must not only be morally right, but must be done for the reasons that make them right. That is to say, the agent's motivating reasons and her moral reasons must coincide.²⁰

Responsiveness to the moral reasons may be taken to require moral understanding of the sort discussed by Hills. The thought is that if one is to act for

¹⁸ See for instance Crisp, "Moral Testimony Pessimism," Hopkins, "What is Wrong with Moral Testimony," Howell, "Google Morals," and Nickel, "Moral Testimony and Its Authority."

¹⁹ Hopkins, "What is Wrong with Moral Testimony," 634.

²⁰ For the canonical expression of this position see Nomy Arpaly, *Unprincipled Virtue* (Oxford: Oxford University Press, 2002). She calls this the 'praiseworthiness as responsiveness to moral reasons thesis.' Julia Markovitz ("Acting for the Right Reasons," *Philosophical Review* 119, 2 (2010): 201-42) defends a similar position under the heading of 'the coincident reasons thesis.'

the reasons that make one's action morally right (and therefore be worthy of moral praise) one must grasp the relation between these reasons and one's action. On this proposal, the moral reason not to defer to the testimony of moral experts is that true moral beliefs or mere moral knowledge acquired in this fashion cannot serve to motivate morally worthy action. If we are responsible moral agents, we are of course concerned with doing the right thing, morally speaking. However, this is not the end of what we are reasonably concerned with, when it comes to our actions. We might have reason to (and perhaps even morally ought to) care also about doing the right thing for the right reasons. Given this other concern, the fact that there are both moral and evidential reasons counting in favour of deferring to experts will not settle the matter. There will be a further question about how these reasons can best be weighed against the moral reasons against such deference. If the pessimist is right, the latter kind of moral reason will at least sometimes win out.

There are problems with this argument for moral testimony pessimism. One is that the account of moral worth from which it derives its force is itself controversial. Moreover, even among those who defend this sort of account there is disagreement about whether it poses a problem for moral deference.²¹ We will set these concerns aside however, and suppose that moral deference would in fact interfere with the performance of morally worthy actions.

Still, this argument has an air of self-indulgence. Morally right actions that are not done out of an awareness of the reasons that make them right are still morally right. They may not attain the ideal of moral worth, but a less than morally ideal right action is nonetheless always morally preferable to a wrong action. If this is not granted we lose all grip on these notions. A pessimist who appeals to the sort of view sketched above thus seems to be offering surprising, if not paradoxical, counsel to someone like Gillian. The pessimist objects to deference to moral experts because a non-expert who does this is thereby cutting herself off (at least locally) from the ideal of morally worthy action. However, the non-expert knows that in not deferring she is increasing the probability that she will perform morally wrong actions. As we pointed out earlier, it is likely that in neglecting to defer to the testimony of moral experts, she is thereby proceeding in a way that is itself morally wrong. What the pessimist must claim, then, is that broadly aretaic considerations to do with the moral worth of one's own actions ought (at least sometimes) to trump the straightforward deontic considerations

²¹ Julia Markovits ("Saints, Heroes, Sages, and Villains," *Philosophical Studies* 158 (2012): 289-311) claims that actions performed because a reliable moral expert says they are right can be morally worthy on her approach.

that in not deferring one is (possibly) already doing something wrong, and (certainly) making oneself more likely to do wrong in the future.

This will be a bitter pill for Gillian to swallow, insofar as she is in fact a morally responsible person. If she has the option of taking steps that will render her less likely to act wrongly, without doing anything morally wrong along the way, it is hard to see how she can be morally justified in not doing so. In tandem with the evidential reasons for deferring, pessimism appears to be both morally and epistemically more dubious than optimism.

We believe this discussion reflects a general issue for pessimists about moral testimony. It is easy to make the following claim look plausible: there is some *pro tanto* reason not to defer to moral experts. This reason might be grounded in epistemic considerations of the kind Hills had in mind, or in moral ones of the kind the current argument is appealing to. However, it is much harder to establish the more interesting claim that *pro tanto* considerations of either sort justify the *sans phrase* claim that it is better not to defer to moral experts in any actual case. On the moral side, we may well rank actions that are morally worthy and morally right more highly than those that are just morally right, but also rank the latter more highly than those that are morally wrong. On the epistemic side, we may well rank moral beliefs that are true and combined with understanding more highly than those that are true but not combined with understanding, but also rank the latter more highly than those that are false. Given that moral deference is conducive to morally right action and true moral belief, then, claims to the effect that reliance on deference cuts one off from something else that is valuable have limited impact. We can call this the ‘*pro tanto* problem’ for pessimism about moral deference.²²

6. No Virtue in Rectitude

There is another line that has made some headway in the literature, one which appeals to moral reasons of a somewhat different sort. On this approach, the problem with moral deference is that this practice interferes with the development of a morally virtuous character. Thus Robert J. Howell writes that while

[t]here might be epistemic dangers associated with moral deference... the real harm is the crippling effect such deference can have on the moral character of the deferring agents.²³

²² See Enoch, “A Defense Of Moral Deference” for a forceful expression of this problem.

²³ Howell, “Google Morals,” 412.

On this view, again, there are moral reasons against forming one's moral beliefs on the basis of expert testimony, only now these reasons are taken to bear on the character of the agent himself as opposed to his action. And again, the claim that reliance on moral experts can serve as an obstacle to the development of a morally virtuous character has some plausibility.

The *pro tanto* problem rears its head again here, however. Let us start once more by considering only the moral reasons on both sides. The moral reasons that appear to support the pessimist argument are – also again – self-regarding, rather than other-regarding. They bear on one's own moral character, and the moral reasons one (presumably) has to promote it. It is doubtful that one can make a plausible moral case against moral deference on such grounds. Consider that if Gillian opts not to take an expert e-reader with her, she is likely to do wrong in situations where, had she heeded the e-reader's counsel, she would have done right. This does not just have an impact on Gillian herself. Her morally wrong actions can have serious, even fatal consequences, for the other denizens of the island. Even in much less extraordinary circumstances, when we do the morally wrong thing we very often cause unnecessary harm to others. Bearing this in mind, it courts the charge of moral preciousness to say that the non-expert ought not to defer out of concern for her own virtue. Once more, the evidential reasons and the moral reasons counting in favour of deference align.

7. Pessimistic About Pessimism

None of the three arguments in support of pessimism about moral testimony discussed above strikes us as attractive. They at best establish that there is some *pro tanto* reason not to defer, which does not itself justify pessimism *sans phrase* about moral deference, either in general or in any particular case. This is the *pro tanto* problem. The failure of the three surveyed arguments for pessimism does not show that there is no good argument to be made.

What it does suggest is that the moral testimony pessimist will have to appeal to a different kind of pragmatic reason than was countenanced above. One possibility worth exploring is that there is something intrinsically morally wrong with moral deference itself. Perhaps there is a strong moral duty not to form one's moral beliefs in certain ways. This latter sort of position has not been explored carefully in the literature. Whether an argument for this kind of view can avoid the danger of high-handedness remains to be seen. As things stand at present in the literature, we are pessimistic about pessimism.

8. In the Pragmatic Mirror

Because of the particular difficulty with finding strong epistemic reasons in favour of *sans phrase* pessimism, the disagreement between moral testimony optimists and pessimists depends to a significant extent on what kind of pragmatic reasons for belief one takes there to be. It is therefore a particularly interesting feature of the moral deference debate that the moral reasons for deferring, or not deferring, track to a significant degree a more general distinction between two types of pragmatic reasons for belief.

There are pragmatic reasons for belief that are consistent with alethic norms. Call these pragmatic reasons ‘convergent.’ The most straightforward examples of convergent reasons occur when what might be called ‘leaps of knowledge’ are possible.²⁴ Leaps of knowledge cases are ones in which any of the relevant beliefs that one comes to have will be true, on account of having it. The classic example of this is Gilbert Harman’s power of positive thinking example (slightly modified here).²⁵

Suppose that Larry has an illness that is, through some mechanism, connected to his doxastic states about the illness. Larry learns about this illness from his doctor, who tells him that, if he believes he will recover, then that belief will in fact cause him to recover. However, if Larry either does not believe that he will recover, or believes that he will not recover, then he will not recover.

On being informed of his illness and its relation to his doxastic states (and assuming that those states are luminous to him), Larry will either have sufficient epistemic reason to believe that he will get better, or he will have sufficient epistemic reason to believe that he will not get better. He will have sufficient reason for the former state if he believes that he will get better, and he will have sufficient reason for the latter state if he either does not believe that he will get better, or believes he will not get better.

Suppose Larry, when he receives the news from his doctor, reasons that because the illness is news to him and he therefore has no particular belief about whether he will recover, he will not recover (having lacked an antecedent belief that he will recover). In doing so, he has arrived at a belief for which there is sufficient epistemic reason. Despite having settled into a stable doxastic state, Larry has strong pragmatic incentives to switch to believing that he will recover. If he does so, there will be no alethic cost - he is guaranteed still to have a true

²⁴ See Reisner, “Leaps of Knowledge.”

²⁵ Gilbert Harman, “Rationality,” in his *Reasoning, Meaning, and Mind* (Oxford: Oxford University Press, 1999), 9-46.

belief about his condition and prospects for recovery – and it is pragmatically better for him to believe that he will get better.

There are other types of cases that take a similar form.²⁶ What all these cases have in common is that they provide pragmatic reasons for belief that do not violate a general alethic constraint on reasons for belief. Doing well morally or prudentially with respect to one's beliefs only requires one to believe what one knows to be true, conditional on and because of one's believing it.

The other general type of pragmatic reasons for belief are those that run contrary to the truth and/or to the agent's total evidence. Call these 'non-convergent.' Non-convergent reasons take on a variety of forms, but there have been three general proposals about how they might arise.

One way in which such reasons might arise is from certain kinds of constitutive norms. Sarah Stroud has argued, for example, that it is a constitutive feature of friendship that one is epistemically partial.²⁷ Evidence that one's friends are dishonest or disloyal must receive less weight than the same evidence that those with whom one is not friends are dishonest or disloyal must be given. While Stroud's particular claims involve friendship, the general line of argument is in principle extendable to participation in other kinds of relations or institutions.

Several authors have offered various examples in which there are strong moral or prudential incentives to believe against the evidence.²⁸ This is the second way in which non-convergent reasons may arise. One type of incentive driven reasons arises from external threats or inducements. A sufficiently knowledgeable and powerful being can provide moral incentives to believe against the evidence. These examples take the following general form: the mad scientist will do *x* valuable thing, if you believe *y*, even though the evidence suggests that *y* is not the case.

A separate, but very interesting, class of cases is due to Berislav Marušić. These cases rely on the intertwining of the belief norm on intending and promising with situations in which one would have to believe in the face of evidence against the required belief.²⁹

²⁶ For extended discussion see Reisner, "Leaps of Knowledge," as well as Reisner, "A Short Refutation."

²⁷ Stroud, "Epistemic Partiality."

²⁸ See Sven Danielsson and Jonas Olson, "Brentano and the Buck-Passers," *Mind* 116, 463 (2007): 511-22, Reisner, "The Possibility of Pragmatic Reasons," and Stephen Stich, *The Fragmentation of Reason: Preface to a Pragmatic Theory of Cognition* (Cambridge: M.I.T Press, 1993).

²⁹ Berislav Marušić, "Promising against the Evidence," *Ethics* 123, 2 (2013): 292-317.

It will be helpful to look at two examples. The first is loosely adapted from Marušić's own. It is *The Sociologist's Marriage*. We may imagine that two sociologists are at the altar, about to pronounce their marriage vows to each other. The presiding official at the wedding asks them to follow up various promises with 'forever and ever.' Each sociologist knows that there is less than a one in two chance that they will remain married, or even in love, for the remainder of their natural lives - or even for more than a decade. If there is a belief condition on promising that requires one not to believe that one's promise will not be upheld, then the sociologists can only sincerely take their vows if they believe against the evidence.

The second example exploits a similar belief condition on intention: that one can intend to do something only if one does not believe that one will not do it. This condition makes intending to perform actions with low chances of success problematic. One may imagine that Shackleton, on his small boat en route to Elephant Island, intended the entire way to rescue his shipwrecked crew. We can assume that having that intention was central to his prospects for success, in focusing his activities, providing confidence to his men, etc. Having this intention, if we accept the belief condition, required his not forming the belief best supported by the evidence, namely that he would drown and die a horrible death in the violent southern seas. In both the Sociologist's Wedding and the Shackleton examples, we have cases in which there are prudential and moral inducements to make promises and form intentions that require believing against the evidence. The pragmatic reasons for believing against the evidence in these cases are parasitic on the ordinary moral reasons for promising and intending, rather than arising from direct incentives, as in mad scientist examples.

Returning to Gillian's Island, it is clear that the moral reasons for belief that tell in favour of optimism are convergent reasons. They are not leaps of knowledge cases, but they share the important feature that the truth of the beliefs and the goodness of having them is non-accidentally connected. On Gillian's Island, Gillian will be forced to make morally consequential decisions, the likes of which she has never before faced. If we grant the assumptions of the case, that Gillian is enkratic and that acting on the basis of true moral beliefs more often yields morally better results than does acting on false moral beliefs, then she will have convergent moral reasons for forming moral beliefs based on the expert's testimony.

On the other hand, it is a clear consequence of rejecting moral testimony that the number of Gillian's true, situationally relevant moral beliefs will be much lower than if she were to accept it. If there are moral reasons for requiring moral

understanding or cultivating moral virtue, at the expense of having true moral beliefs or mere moral knowledge, then those reasons are non-convergent.

Conclusion

We have argued that the case for moral testimony pessimism does not look promising at present, either epistemically or morally. This is not to say that there are no *pro tanto* reasons, either epistemic or moral, for pessimism. Rather, it is difficult, at least for us, to see how they will add up to a defence of all-things-considered pessimism.

We have, in effect, suggested that the implicit rankings of epistemic and moral goods put forward by various pessimists are incorrect. We expect that pessimists will disagree. It is thus interesting to consider briefly what it would take for a pessimist to offer a convincing argument in favour of her view.

With respect to epistemic reasons against moral expert testimony, we are not convinced that there is a case to be made. If one denies that there are moral experts, or that they are identifiable, then the issue will be moot. However, unless one is willing to make the rather strong claim that it is conceptually or metaphysically impossible that there are moral experts, the matter's being practically moot does not settle the theoretical question.

That leaves the pessimist with the burden of explaining why acquiring moral beliefs in the same way that one acquires most of one's other beliefs is epistemically problematic. One port of call might be to object to testimony about the *a priori*, but with a little reflection this makes not a small part of most people's knowledge of maths and also sciences that are partially mathematical deeply problematic. It is up to the pessimist to show how we can avoid throwing the baby out with the bathwater, or why we ought not to mind doing so. We are sceptical that the pessimist's epistemic burden can be discharged.

If that is right, then the pessimist must defend pessimism by appealing to pragmatic reasons for belief. In principle, there is nothing objectionable about doing so. We have argued that the pragmatic reasons for belief appear to tell in favour of the optimist. To respond convincingly to our criticisms, the pessimist will have to make a case relying on non-convergent pragmatic reasons for belief. While we certainly do not object to the view that there are non-convergent reasons, it is nonetheless interesting to note that pessimists are committed to the existence of the most radical kind of pragmatic reasons for belief, those which are contrary to evidence and often to truth.

Moreover, if all-things-considered pessimism is to be justified, these non-convergent reasons must be weighty enough to override the combined force of the

Andrew Reisner and Joseph Van Weelden

moral and evidential considerations on the optimist's side. While nothing we have said strictly rules out *sans phrase* pessimism, we see little reason at present to think that there are more than *pro tanto* reasons, both epistemic and pragmatic, for rejecting moral testimony.

A CASE FOR EPISTEMIC AGENCY

Dustin OLSON

ABSTRACT: This paper attempts to answer two questions: What is *epistemic agency*? And what are the motivations for having this concept? In response to the first question, it is argued that epistemic agency is the agency one has over one's belief-forming practices, or doxastic dispositions, which can directly affect the way one forms a belief and indirectly affect the beliefs one forms. In response to the second question, it is suggested that the above conception of epistemic agency is either implicitly endorsed by those theorists sympathetic to epistemic normativity or, at minimum, this conception can make sense of the legitimacy of the normative notions applicable to how and what one should believe. It is further contended that belief formation in some respects is a skill that can be intentionally developed and refined. Accepting this contention and the existence of certain epistemic norms provide inconclusive yet good reasons to endorse this concept. Recent challenges to this concept by Hillary Kornblith and Kristoffer Ahlstrom-Vij are also considered.

KEYWORDS: belief formation, belief control, epistemic agency, epistemic normativity

Introduction¹

In this paper I respond to two questions: What is *epistemic agency*? And what are the motivations for having this concept? I argue that epistemic agency (EA) is the agency one has over one's belief-forming practices, which will directly affect the way in which one forms belief and indirectly affects the beliefs one forms. There are a number of reasons to adopt this concept. First, it is implied in extant theories of epistemic normativity. For example, if one accepts that there are epistemic virtues or epistemic duties, then one is implicitly accepting some notion of EA. A second reason to adopt EA is that belief formation is in some sense a skill; we can improve as believers. Doxastic self-improvement and skilled belief formation are marks of EA. I suggest, then, that there are inconclusive yet good reasons to endorse this concept.

Recently some have registered skepticism about the viability of EA. Hilary Kornblith, e.g., regards EA as mythological because we do not control the actual

¹ Thank you to audience members at the University of Rochester and at Cornell University's Ethics of Belief Workshop for helpful comments and questions. And special thanks to the two anonymous referees, Patrick Bondy, Richard Feldman, Kate Nolfi, Sharon Ryan, and John Turri for helpful comments and discussions.

reason-responsive mechanistic processes that bring about a belief.² If we cannot control that aspect of belief formation, he argues, we are not agents with regards to belief formation. A second challenge to EA is offered by Kristoffer Ahlstrom-Vij.³ He argues that even if we can give an account of EA that avoids Kornblith's challenge by appealing to higher-order processes like reflective reasoning and deliberation, there are reasons to reject this concept. Ahlstrom-Vij contends that a key motivation for having a concept like EA is that we can practice epistemic self-improvement; there are important ways that we cannot self-improve epistemically, however. Because of this limitation, we cannot then be said to have EA. I agree with Kornblith that we do not have direct control over the lower-level mechanistic components of belief formation. I also agree with Ahlstrom-Vij that epistemic self-improvement is one of the key motivations for developing and accepting a concept like EA. I disagree, however, with both of their conclusions. This paper proceeds, then, as follows. First I present and develop my conception of EA. I next make a case for why we should accept this conception. And finally, I address and reject the aforementioned arguments against EA.

1. Epistemic Agency

I propose that epistemic agency is distinct from other notions of belief control. It does not require one's ability to choose one's belief or one's ability to decide to believe – e.g. I want to believe that p , so I believe that p . Nor is it a concept that simply requires one's ability to control or influence the formation of specific beliefs – e.g. by A -ing I will come to believe that p .⁴ Rather, I contend that epistemic agency denotes the motivation and ability to refine and alter one's belief-forming methods and subsequent belief-forming practices – these methods and practices can collectively be understood as a *doxastic disposition* or belief-forming abilities, which can also be characterized as one's propensity to form true or false or coarse-grained or fine-grained beliefs within different domains. Epistemic agents have the ability to hone and refine their belief-forming abilities through altering their doxastic dispositions.⁵ In so doing, one can affect one's beliefs, but neither directly nor specifically.

² Hilary Kornblith, *On Reflection* (Oxford: Oxford University Press, 2012).

³ Kristoffer Ahlstrom-Vij, "Why We Cannot Rely on Ourselves for Epistemic Improvement," *Philosophical Issues* 23, 1 (2013): 276-296.

⁴ This process is certainly something an agent does. But, as will be clarified below, I distinguish this type of voluntary belief control from EA.

⁵ Recent works on belief control has made appeals to dispositional characteristics as playing, or potentially playing, an active role in the beliefs or judgments we make. Kate Nolfi, "Why is

The above suggests that EA is distinct from other belief-control concepts. The concepts I have in mind fall under the heading of doxastic voluntarism, which comes in two forms: direct and indirect. *Indirect doxastic voluntarism* (IDV) is the thesis that we can form beliefs through indirect voluntary control. For example, I want to form the belief that the lights are on in my office, so I go to my office and turn the lights on. As long as my perceptual apparatus is working as it should, that I turn on the lights and observe their illumination will result in my believing that the lights in my office are on.⁶ *Direct doxastic voluntarism* (DDV) is the thesis that we can form our beliefs through direct voluntary control. For example, if one were able to believe that p at will – decide to believe that p – then DDV would be true.⁷ Closely related to DDV is *doxastic freedom*. For one to have doxastic freedom is for one to have the freedom to exert direct voluntary control over one's doxastic attitudes.⁸

I argue that EA is not reducible to IDV or DDV, and that EA does not require that the agents in question have doxastic freedom. EA neither reduces to DDV nor requires one to have doxastic freedom. I agree with the majority of philosophers who conclude that both of these theses are implausible. One cannot simply decide to believe; one cannot freely choose one's doxastic attitude. That is, doxastic states are not the types of things that we have direct agential control over.

Epistemic Evaluation Prescriptive?" *Inquiry: An Interdisciplinary Journal of Philosophy* 57, 1, (2014): 97-121, e.g., defends a causal account of doxastic control, wherein one's adoption of how one ought to believe causally effects how we actually believe. One such way this might play out, which Nolfi finds promising but falls short of endorsing, is that our judgment about what we ought to believe exerts "causal power in shaping the ways in which we are disposed to regulate our beliefs" (110). She dubs this disposition to regulate belief as one's *cognitive character*. Conor McHugh, "Epistemic Responsibility and Doxastic Agency," *Philosophical Issues* 23, 1 (2013): 132-157, also appeals to dispositions as way to combine reason responsiveness with agency – or, in his terms, doxastic agency – in his "Reasons-Action Principle: Responding to a putative reason involves conscious agency, at least dispositionally" (146). Where one's disposition is in part the ability to respond to reasons through recognizing them as reasons as such; recognition and responsiveness are not passive and thus require agency. While both of these theories of belief control are not inconsistent with the one presently defended, my proposal is broader.

⁶ This example is from Richard Feldman, "The Ethics of Belief," *Philosophy and Phenomenological Research* 60, 3 (2000): 671-672.

⁷ This view is not widely accepted, although it does have defenders. See, e.g., John Turri, David Rose, and Wesley Buckwalter, "Choosing and Refusing: Doxastic Voluntarism and Folk Psychology," unpublished, Matthias Steup, "Doxastic Freedom," *Synthese* 161 (2008): 375-392, and Patrick Bondy, "Epistemic Deontology and Strong Doxastic Voluntarism: A Defense," *Dialogue* (2015), DOI: <http://dx.doi.org/10.1017/S0012217315000487>.

⁸ A defense of doxastic freedom can be found in Steup, "Doxastic Freedom."

The question then becomes: where is the agential control that I am suggesting we have in the formation of a belief?

The control needed for EA is more closely related to IDV than DDV. I am not proposing that EA is simply a form of IDV, however. In IDV, one acts with the goal to form a specific belief, as in flicking the light switch so as to form the belief that the light is on. EA does not require that a specific belief be formed. One does not exercise one's epistemic agency with the goal of forming a specific belief; when one exercises one's epistemic agency one has the goal of forming beliefs in a specific way - e.g. to maximize truth while avoiding falsity or to form more fine-grained beliefs given a specific body of evidence. EA is similar to IDV, however, in that IDV merely requires normal agency for the desired belief to be formed. Likewise, to adjust one's doxastic disposition, one merely needs agency as we commonly understand it.

Consider, for example, candidate higher-level processes involved in belief formation, processes like deliberation, reflection, attentiveness, and intentions. These processes are the types of things we can have control over. High-level processes coupled with the lower-level mechanistic reason-responsive processes of coming to hold or not hold certain beliefs, then, exemplify EA. That is, we can control certain processes as a means to altering how we believe.⁹ Reflection and deliberation, e.g., are types of practices that can influence belief formation and are things we can control.¹⁰ The role that these processes can have in EA is through one's attempts to achieve one's epistemic goal through reflecting and deliberating on evidence or through what one has accepted as sufficient or insufficient evidence or through what one accepts as an epistemically virtuous principle by which to achieve one's epistemic goals. This reflection does not facilitate one's choosing what to believe; rather, it informs one's belief-forming process for *how* to believe given one's epistemic goals – *viz.* purposefully undertaking steps that will alter our doxastic dispositions. Let us turn to an example.

⁹ Recall that my proposal does not say that we act to form beliefs about specific propositions. In this way I would distance my view from something like Pascal's proposal, as he suggests that we can act so as to believe that God exists and thereby come to believe the specific proposition "God exists." I do not think that we can act to believe in this way. I suspect the only plausible ways in which we can act to believe are uninteresting IDV cases.

¹⁰ There may be one way to slant this so that it seems as though my proposal requires that we do seek out belief in specific propositions. If our world is consistent with classical logic, then if one were to have the goal to form only true beliefs, one's forming a true belief about a specific proposition could be seen as one attempting to believe that exact proposition because it is the only proposition that would satisfy the inquiry in question. That is, one would have intended to form that belief in virtue of its being true.

Consider Campbell's desire to form accurate beliefs about scotch so as to be able to converse with other aficionados.¹¹ Currently, Campbell's beliefs are not fine grained concerning the different qualities he experiences when drinking scotch. His desire to form these more fine-grained beliefs motivates him to undertake certain acts that will refine his palate – e.g. he reads the taster's notes for each scotch he drinks, familiarizes himself with the jargon associated with the different scotch-drinking steps, and is phenomenologically sensitive to the experiences he has while drinking scotch. Eventually Campbell's beliefs concerning the different scotches he samples become fine grained. Thus, rather than simply forming coarse-grained beliefs, like merely having the ability to distinguish between a scotch whisky and a rye whiskey, Campbell can now form more fine-grained beliefs, such as the age, region, and casking processes of the different scotches he samples. Observe, then, that Campbell neither chooses to believe any specific proposition, nor exercises doxastic freedom in order to implement the agency associated with EA. The agency Campbell exercises is the straightforward agency we associate with action planning and goal satisfaction. The present example is just a warm up as to how we might think about intentionally altering our doxastic disposition so as to alter the way we form beliefs. More details emerge below.

2. The Case for Epistemic Agency

In what follows I make the case for EA. In making this case, I first discuss how EA can make sense of epistemically normative notions like epistemic virtues and epistemic duties. Those who reject epistemic normativity will not find this defence convincing; those who accept epistemic normativity will already in some sense endorse my proposal. There are additional reasons to accept EA, however. Some ways in which we form beliefs can be understood as a skill, an ability that can be developed – i.e. we can intentionally attempt doxastic self-improvement and achieve our epistemic goals. As forming beliefs can be a developed ability, it requires some type of control – the type I suggest is found in my formulation of EA. So, even if one remains unconvinced regarding epistemic norms, the fact that we can set and achieve some epistemic goals, i.e., we can undertake steps for doxastic self-improvement and become more skilled epistemically, is further reason to adopt EA.

¹¹ This implies that practical interests can motivate some of our epistemic interests. Some may find this difficult to accept. I don't.

2.1 Epistemic Agency and Epistemic Normativity

One reason to accept the current portrayal of EA is that it can help us make sense of different normative notions in epistemology. It is plausible that we can form goals or have motivations for how we come to form beliefs or the practices we use to form true beliefs. There is also reason to think that many of our candidate theories of justification have a motivational element. Consider two normative epistemic ideas: epistemic virtues and epistemic duties. I propose that if we accept, even if only in part, that there are epistemic virtues or epistemic duties, then some type of EA is required. One reason to accept EA, then, is that it is already implied in a number of theories.

EA and Epistemic Virtues

It is clear that much of virtue epistemology takes for granted the type of agential process that I am here espousing. According to a number of virtue epistemologists, cognitive or epistemic virtues have a motivational component. These virtues are habits developed with the motivation for knowledge, which in virtue terms is true belief that results from a cognitive virtue. Regardless of whether we accept this definition of knowledge, there do seem to be epistemic virtues, and they do seem to be motivated by the desire to acquire true beliefs. Thus, part of the appeal of cognitive virtues is the motivational component to become a more virtuous believer. This motivational component is analogous to the motivational aspect of EA: that one desires - has some motivation - to refine or improve one's belief-forming habits, in general or over a specific domain. Compare this proposal with Linda Zagzebski's description of the motivational component of a cognitive virtue:

The primary motivation underlying the intellectual virtues is the motivation for knowledge. Such a motivation clearly includes the desire to have true beliefs and to avoid false ones, and... such a motivation leads a person to follow rules or procedures of belief formation that are to her epistemic community to be truth conducive.¹²

On this view, then, if it we can acquire epistemic virtues, then we must have the ability to develop these cognitive habits. Or consider John Greco's characterization of a cognitive virtue, which again highlights the importance of the agent's role in acquiring and fostering a virtue:

A virtue, in one important sense, is an ability. An ability, in turn, is a stable disposition to achieve certain results under certain conditions. Further, when we

¹² Linda Zagzebski, "Virtues of the Mind (selections)," in *Epistemology: An Anthology*, ed. Ernest Sosa and Jaegwon Kim (Blackwell Publishing, 2000), 465.

say that a subject *S* has an ability to achieve certain results, we imply that it is no accident that *S* achieves those results. *S*'s disposition to achieve the relevant results is grounded in certain properties of *S*, such that under the appropriate conditions any subject with those properties would tend to achieve those results.¹³

Again, it seems clear that EA captures the idea that we have an ability to achieve certain doxastic goals via developing one's doxastic disposition – or as the virtue theorists term it, to form cognitively virtuous habits.¹⁴

EA and Epistemic Duties

Epistemic deontologism holds that there are things we epistemically *ought to* or *ought not* believe.¹⁵ That is, we have epistemic duties to believe in a certain way. Much like the motivational component of epistemic virtues, epistemic duties also have a motivational component. Here, however, the epistemic goal is to form and maintain appropriate beliefs, according to the duty one has to believe in the appropriate way. That is, one ought to believe that *p* when it is epistemically appropriate to do so. From the evidentialist perspective, for example, if one's evidence supports that *p*, then one ought to believe that *p*; if one's evidence does not support that *p*, then one ought not believe that *p*. Thus, if one is not suitably attuned to one's evidence, then one is falling short of one's epistemic duty.¹⁶ If we have EA, however, then we can develop our abilities such that we do believe in the right way, *viz.* we form beliefs that are sensitive/responsive to the right evidence.

¹³ John Greco, "Virtue and Vices of Virtues Epistemology," in *Epistemology: An Anthology*, ed. Ernest Sosa and Jaegwon Kim (Blackwell Publishers, 2000), 468.

¹⁴ I am not defending Virtue Epistemology. I am merely attempting to show that if there are epistemic virtues – i.e. a motivation and habitual ability to form true beliefs and avoid false ones – then they are best understood in terms of our having EA.

¹⁵ The issues involved in deontological epistemology are complicated. Sorting through these issues would take us well beyond the scope of this paper. For a discussion concerning some of these issues see Feldman, "Epistemic Deontologism." I merely here wish to highlight that if one prefers to think of epistemic justification in deontological terms, then EA is for that person.

¹⁶ The *ought* being proposed here is epistemic; epistemic deontologists are concerned with what one epistemically ought to believe. This *ought* is distinct from other types of *ought*, like moral or prudential oughts, although some have made the case that the two are more closely related than they might seem. The *locus classicus* here is William K. Clifford, "The Ethics of Belief," in *The Theory of Knowledge*, 3rd edition, ed. Louis Pojman (Canada: Wadsworth, 2003), 515 - 518. For a contemporary defense of Clifford's proposal see Sharon Ryan, "In Defense of Moral Evidentialism," *Logos & Episteme. An International Journal of Epistemology* VI, 4 (2015): 405-427.

A common complaint about epistemic deontology is that it seems to require a level of voluntary control over our beliefs that we do not have. The argument runs as follows:

1. If we have a duty to x , then we can be held responsible for whether or not x .
2. We cannot be held responsible for something we do not have voluntary control over.
3. We do not have voluntary control over our beliefs.
4. Therefore we cannot be held responsible for whether or not we believe that p
5. Therefore we do not have epistemic duties.

In response to this argument, I suggest that defenders of epistemic deontology should, or implicitly do, accept my version of EA. EA provides a way to avoid the above criticism without having to deny the quite plausible first premise. Epistemic deontologists can reject premise (2). We can be held responsible for our beliefs because there are things we can do that can affect them - we therefore do have the right kind of control to allow for epistemic duties.

The kind of control that we have over our beliefs is indirect. Furthermore, this control is voluntary only in the sense that we can undertake certain practices of our own volition; it is not the type of voluntary control found in (2). The type of voluntary control being referred to in (2) suggests that voluntary control occurs over the formation of specific beliefs by an act of will – *viz.* DDV. This is not the type of control being defended in EA. Regardless, the type of control being defended in EA is sufficient for responsibility.

Take one who has a skewed view of the evidence for a certain set of beliefs. Suppose Jones, e.g., believes that the university is conspiring against his daughter, which explains why she is doing poorly in her classes. In Jones's mind, his daughter's lack of success is not due to an absence of effort or skill on her part, but rather the result of a conspiracy between faculty and administration to ruin the student's academic career. The evidence that Jones uses to justify these beliefs is based on grade reports and instructor comments. However, he also has competing evidence: he attended orientation; he observed all of the student-success-based programs offered by the university – e.g. a well-stocked library with lots of quiet study space, subject tutors, mental health specialists, and so on; and he also attended college and is thus able to appreciate the rigors of adjusting to freshman year. If Jones were to be more impartial, he would further recognize that it is against the university's interests to have students do poorly. Because the evidence

is such that the father's conclusion is clearly false and because he himself has sufficient evidence supporting the falsity of his belief, we can conclude that he shouldn't believe that the university is conspiring against his daughter. His judgment of the evidence is at once incorrect and something he could change with a more disinterested reflection. This is a case where the believer has all the evidence needed for the correct belief, yet forms the wrong conclusion by unduly giving too much weight to an inconclusive or defeated subset of available evidence. He could and should believe otherwise.

We can also think of cases where a belief is formed too hastily on insufficient evidence or a belief is not formed when it should have been due to one's having sufficient evidence. Suppose Black visits Germany and happens to see a black cow. She infers from this observation that all the cows in Germany are black. This conclusion seems too hasty. Although there is some evidence supporting it, Black should not hold this belief because her evidence is seemingly insufficient. Conversely, suppose Green is an agricultural expert. She has experienced many European cow breeds. She has observed that Germany has almost exclusively Holsteins, which have a spotted coloring.¹⁷ She is quite skeptical however, and does not conclude that most of Germany's cows are spotted, even though she has epistemically sufficient reasons to form this belief. Green should believe that most German cows have a spotted coloring.

Being more objective or disinterested in how one interprets one's evidence upon reflection would prevent Jones forming the false belief about his daughter's university. Adopting a more skeptical attitude would prevent Black from forming beliefs too hastily and becoming less skeptical will prevent Green from unwarranted agnosticism. These dispositional characteristics are the types of things we can alter over time, indirectly affecting the beliefs that we form.¹⁸ We can voluntarily take steps to become more attuned to what the existing evidence is, to be more sensitive to what that evidence supports, and more apt to form the appropriate doxastic attitude given that evidence without having to voluntarily believe a specific proposition. These steps are facilitated by EA. By accepting EA, defenders of epistemic deontology are able to avoid this standard criticism made against their view.

¹⁷ Germany in fact has over 40 breeds of cow, but for the sake of the example we can assume that it has almost exclusively Black Holsteins.

¹⁸ How we can alter these doxastic dispositions and the types of steps we can take in so doing were exemplified in §1, with Campbell's goal to form more refined beliefs about scotch. This type of process is more fully discussed in §2.2 below.

It might be countered that although it is plausible that we can do things to affect how we form beliefs, which may allow for the possibility of epistemic duties, it is strange to praise or blame someone for her belief. As Richard Feldman observes, “praising someone by saying something like, ‘That was a really great bit of believing you did there’ sounds bizarre.”¹⁹ And while it does sound bizarre to praise or blame someone in such a direct manner for what one believes, there are other similar ways in which the praise or blame is implicit. It is not nearly as bizarre sounding to hear something like, “I sure am glad that Judy is on the murder case; she can get to the truth like no other” or “John sure knows how to pick ‘em!” implying that John is a poor judge of character and is bad at forming beliefs about those with whom he associates. We do assess the beliefs of others. That believers have EA allows for such assessments. So although we can recognize that baldly asserting praise for a belief sounds bizarre, there are ways of attributing responsibility and assessing one’s beliefs and the ways in which one forms one’s beliefs that are not so bizarre. I suggest that the appropriateness of this latter possibility is due to EA.

It appears, then, that if we accept some form of epistemic normativity, whether by accepting that there are ways we should believe or ways in which we should form beliefs, we are accepting some form of agency. Accepting that satisfying these motivations does not require the ability to have direct voluntary control over our beliefs, there thus seems to be some plausibility for EA as developed here.

2.2. Doxastic Self Improvement and Skilled Belief

Not all will be convinced by epistemic normativity, however. Fortunately, there is another reason to accept the conception of EA here defended: belief formation is a skills-based endeavor. Some people are better than others at forming beliefs in different areas and in different ways. This observation suggests, then, that some beliefs can be formed skillfully. *A fortiori*, belief formation can be viewed in some respects as a skill. If we accept that some people are more skilled, that is, are better than others in forming true or more detailed beliefs in certain domains, and these skills result in part from the higher-level processes discussed above, then one’s belief-forming abilities can be intentionally refined and developed. If one can intentionally do anything, then agency must be involved somewhere in the process.²⁰ One’s ability to alter one’s doxastic disposition enables the capacity for

¹⁹ Feldman, “Epistemic Deontology,” 353.

²⁰ I assume that intentions imply agency. I therefore leave this claim unsupported. A brief discussion of intentions and doing something intentionally is included in §3.2 below.

one to improve one's skills that affect one's belief-forming abilities. In other words, we can undertake doxastic self-improvement in some ways as a result of EA.

On my formulation, while not sufficient for EA, phenomena like deliberation, reflection, and other dispositional characteristics involved in the belief-forming process are paradigm representations of the types of higher-level processes involved in EA. These processes can influence the lower-level processes, that we do not directly control, which then will affect what one comes to believe through how one forms one's beliefs. Take deliberation, for example. Deliberation is something we do. Focusing our attention on different pieces of evidence, seeking further evidence prior to committing to a belief, and further reflecting on reasons to believe can all be included as elements of the deliberative process. Suppose I am a jury member and must form a belief about the defendant's guilt or innocence. I do not passively accept all of the information that is presented to me during the trial. I deliberate and weigh the evidence presented; I play a more active role in the belief-forming process. Ultimately a belief will emerge from these deliberations, one that at the lower levels I cannot control, one that is reason responsive. Regardless, this type of indirect influence remains something we do and it can affect how and what we believe. We thus have reason to accept the potential influence of our higher-level processes and the mechanistic aspects of belief formation.

That higher-level processes can influence belief formation speaks to another reason for accepting EA, and also involves defending my second claim from above: that belief acquisition can be skillfully undertaken. We can intentionally improve some of the ways in which we come to hold our beliefs, just as we can refine a number of other characteristics. A carpenter's ability to hammer nails into boards with one swing is a skilled movement. By the simple fact that the carpenter's job involves hammering nails often, she improves this skill over time. Suppose, however, that a non-carpenter wishes to strike nails with similar skill and efficiency. What should the non-carpenter do to achieve this goal? The obvious response is that the non-carpenter do what the carpenter has done, namely, hammer a lot of nails. The non-carpenter can intentionally undertake the process of acting like the carpenter to eventually become a skilled nail-striker.

I submit that honing our abilities as belief formers in certain areas is similar to refining one's ability to strike nails. The carpenter's ability to hammer the nail was not something consciously acquired; the ability emerged over time as a result of hammering nails. Most of us form many of our beliefs in ways similar to the carpenter hammering nails. We simply form beliefs as part of our role as belief

formers. In some cases we acquire more skill and refinement in forming beliefs simply because of where our attention is most often focused – e.g. generally a judge will be able to form more accurate and refined beliefs about a person’s character than could a plumber, whereas a plumber will form more accurate and refined beliefs about piping than a judge. And, like the apprentice who desires to improve his skill as a nail-striker, so too can we develop skills as believers by using similar methods of imitation and emulation. If Jay wants to form accurate beliefs in a given field, Jay can emulate the experts within that field, practicing their belief-forming habits, and thus improve her own belief-forming habits. In some cases this improvement may simply be a matter of recognizing and gathering the relevant evidence. In other cases the improvement might concern the deliberative process, where no further evidence is needed; one’s ability to judge one’s extant evidence has simply become more fine-grained.

If improvement in nail striking is found in the repetitive emulation of the journeyman carpenter, how is this type of practice possible for belief formers? Recall the case of Scotch-drinking Campbell. Campbell’s attempt at doxastic self-improvement is to emulate expert scotch drinkers by familiarizing himself with the lingo and by extensive epistemic research – i.e. self-reflectively drinking a bunch of different scotches. We observe that this process is a combination of acquiring more information and also reinterpreting the existing information one has. The new information includes, e.g., the scotch-drinkers’ lingo, the different steps included in sampling scotch (nose, palate, and finish), and the nuances found in different scotch-distilling regions. The extension of EA is not merely acquiring new information, however. Consider the existing information – *viz.* the scotch itself. In addition to the new information presented above, Campbell becomes more phenomenologically aware of the extant experiences. By refining his ability to experience scotch tasting, Campbell refines the judgments and subsequently the beliefs that are formed on the basis of these judgments.

We can extend this type of practice to more abstract belief-forming practices. Suppose I want to form more accurate and refined beliefs about human character. How do I imitate the practices and habits of a proven judge? Suppose Judy, a proven judge, gathers evidence by studying the cues she takes from testimony, body language, and the plausibility of certain cases as described by the person whose character is being assessed. Further, after she has collected this evidence, she then processes it by reflecting on it, comparing it to other cases, and by training herself to do so disinterestedly and with proper proportions. If one were to adopt these tactics, over time one would improve one’s abilities as a judge of character. Here, the development of one’s doxastic disposition does not require

the gathering of more information; rather, the development simply requires one to be able to better understand the information one already has access to, like testimony and body language. Certain traits such as disinterestedness, patience in adopting a belief, and properly proportioning the weight given to certain types of evidence would each be aspects of our doxastic dispositions that can be, to some extent, under our control.²¹

The previous two examples require an extended period of time, or broader extension of EA, to alter one's disposition. This needn't be the case, however. More immediate, narrower, applications of EA can occur also. Consider Sophie, a sophomore philosophy student. She is at a social gathering and overhears a senior student have a slip of the tongue, using *ad hoc* instead of *ad hominem*. From this testimony, Sophie forms the belief that *ad hoc* is the fallacy of attacking the person, not the argument. When misusing this phrase, Sophie is corrected by Carrie, who is Sophie's peer. Sophie argues with Carrie about this point but comes to discover that she has formed a faulty belief. Due to this experience, Sophie decides that she should be more judicious with the amount of credence she gives to different types of evidence. She thereby becomes less likely to form rigid conclusions given limited or weak evidence. This way of improving our doxastic disposition can have a nearly immediate effect. Second-year Sophie's judgments will be much less hasty from this point forward.²²

What can we take away from the above examples? First, we observe that the seasoned judge who forms accurate beliefs about character and the seasoned scotch drinker who forms fine-grained beliefs about whisky are like the carpenter who has developed nailing skills as a result of her job. Likewise, the individual who intentionally approaches character judgment like the seasoned judge or who intentionally refines his palate for the nuances of scotch, so as to form more

²¹ Contrast this example with Jones the suspicious father example from §2.1 above. Jones believes a false proposition on the basis of poor evidence selection. Were he to be more objective towards the situation more generally and less hasty to form his judgement that the university is 'out to get' his daughter, Jones would not have formed this belief. These types of practices are the types that we can influence.

²² John Turri, in conversation, has suggested that this type of decision to not be so open to weak testimonial evidence is sufficient for DDV. Turri suggests that the adoption of this attitude is done directly and has an immediate impact on what one believes. This decision and immediate impact, according to Turri, is analogous to deciding to believe. In other words, my proposal might be construed as a type of direct voluntarism. This is a conclusion Turri would endorse as a defender of DDV. I am not a proponent of DDV and do not wish to take on the burden of defending that view. So if Turri is correct, and my arguments work, then so much the better for defenders of DDV. I do not conceive of EA in this way, however. And I doubt that any of those appealing to this concept would accept DDV either.

accurate or fine-grained beliefs, are akin to the apprentice who intentionally emulates the practices of the carpenter to become better at striking nails. And while many of our skills as belief-forming agents are developed unconsciously, that we can consciously undertake certain processes to become more skilled at forming beliefs - *viz.* intentionally emulating those whose skills are already developed in the desired area - provides us with reason to accept EA.

3. Arguments Against Epistemic Agency

I next consider two challenges to EA. The first is Hillary Kornblith's appeal to the mechanistic lower-level process of belief formation that we cannot directly control. The second is Kristoffer Ahlstrom-Vij's contention that we are unable to epistemically improve in important ways.

3.1 Kornblith and the Mechanism of Belief-Formation

Recently Kornblith has taken issue with the notion of reflection as providing the philosophical fruit that some believe it bears. EA is one of the concepts subject to scrutiny within this grander project. Kornblith is never really clear on what he means by *epistemic agency*, i.e. the view that he is attacking, but seems to conclude that whatever else it amounts to, reflection is one of its essential components. In other words, reflection somehow captures what it means to be an epistemic agent because it enables us to affect our beliefs. In other words, reflection somehow captures what it means to be an epistemic agent on Kornblith's understanding of this concept. He contends, however, that higher-level processes, like reflection, are not marks of EA because none of these processes involve directly committing to or altering a belief - appealing to them does not warrant adopting EA. Indeed, on his account, EA is a mythological concept that should be done away with.

Kornblith proposes that when we consider how it is that we come to acquire beliefs, we do not consider ourselves as actively involved in the formation of the belief itself. However, we can be tempted to think we are epistemic agents when

[w]e consciously entertain alternative views, and we think about which, if any, belief about the situation before us we are justified in holding. In situations like this, we seem to play a more active role. We don't just find ourselves believing things. Rather, we decide what to believe; we make up our minds; we choose to believe one thing rather than another.²³

²³ Kornblith, *On Reflection*, 85.

The suggestion, then, is that because we have the ability to weigh evidence and consider alternative possible conclusions, we are tempted to judge ourselves as epistemic agents.

Kornblith argues to the contrary, claiming that this type of agency cannot occur without some form of voluntarism:

If [defenders of EA] wish to insist that we are agents with respect to our beliefs, that there is, in short, genuine [EA], then how are we to make sense of this idea if it is not by way of some sort of voluntarism about belief?²⁴

Kornblith assumes therefore that EA, whatever it amounts to, must include some form of voluntarism in belief acquisition. We can summarize his argument as follows:

1. Either beliefs acquired from high-level processes, e.g. reflection or deliberation, are under direct agential control, or the appeals to higher-level processes do not legitimize the notion of epistemic agency.
2. All beliefs, whether formed from high-level processes or unreflectively, are no different at the lower, mechanistic, level of belief acquisition.
3. There is no agential influence on the lower level of belief acquisition.
4. Therefore, appeals to the higher-level processes of belief formation do not legitimize the notion of epistemic agency.

Premise 1 includes the different phenomena that one might appeal to in defending the notion of EA.²⁵ Kornblith considers the possibility that if any one of these concepts is discovered to be under our control, then perhaps EA is not illegitimate. However, in support of premises 2 and 3, Kornblith considers the aforementioned higher-level processes, rejecting each in turn as marks of EA. Thus, we are to conclude that EA is an illegitimate idea.

Kornblith's denial of EA is motivated by his conclusion that, regardless of the control one might have over any of these higher-level processes, we have no control over the mechanistic lower-level component of belief formation. Take reflective belief as a mark of agency.²⁶ Even when an agent reflects, there is no agency in the actual formation of the belief. Likewise for deliberation when

²⁴ Kornblith, *On Reflection*, 85.

²⁵ Indeed, I make such an appeal in §1.

²⁶ Kornblith cites Richard Moran, *Authority and Estrangement* (Princeton: Princeton University Press, 2001), Ernest Sosa, *A Virtue Epistemology* (Oxford: Oxford University Press, 2007), Christine Korsgaard, *The Sources of Normativity* (Cambridge UK: Cambridge University Press, 1996), and Stuart Hampshire, *Freedom of the Individual* (Chatto and Windus, 1965) as each suggesting that we have EA and that reflection is an essential part of it.

judging what to believe or intentionally acting in such a way as to form a certain belief.²⁷ According to Kornblith, a belief is produced by a reason-responsive mechanistic process; a belief is not produced through deliberate choice or from one's wants or desires. The mechanistic response to reasons is in no way under agential control.²⁸ The appeal to higher-order processes over which we can have some control, then, does not conceptually legitimize the notion of EA.

Clarifying his proposal, Kornblith considers an analogue, the 'screening wand' - a tool used at security stops in airports. These wands are used by screening agents and react to certain types of external stimuli. The suggestion, then, is that any higher-order processes involved in belief acquisition are like the process of directing the wand. The actual production of the belief is mechanical in the same way that the wand's reacting to external causes is mechanical. Our role as belief formers is like that of the security agent who can manipulate the wand but cannot directly control its responses to external inputs. The cognitive response to evidence is similar to the screening wand, in that the response is exclusively produced by external stimuli and not by anything the agent does. In other words, just as there is no agency in the actual screening mechanism, neither is there agency in the actual production of a belief.

Response to Kornblith

I suggest we can reject the assumption made for premise 1 of Kornblith's argument - that EA requires our beliefs to be under our direct control. The main concern with this argument, most notably in this premise, is that Kornblith reduces EA to direct voluntarism about belief but doesn't provide any positive reasons to accept this reduction. The concern, then, is that Kornblith unwarrantedly limits agency to those things that we have direct control over. I argue that agency needn't be limited in this way, but rather that agency can be captured by describing what it is

²⁷ As an example of the type of phenomena Kornblith has in mind when referring to intentionally acting in order to form a belief, consider Pascal's suggestion that if we act as though God exists, partaking in worship and the sacraments, we will eventually come to believe in God's existence. See Blaise Pascal, "The Wager," in *The Phenomenon of Religious Faith*, ed. Terrence Reynolds (Pearson Prentice Hall, 2005), 141-144.

²⁸ Kornblith acknowledges that we are not always perfectly responsive to reasons: "When we offer reasons for belief, however confident we may be that we are in the right, we do not just assume that our interlocutors will come to share our views" (130). But, like any other mechanism, the reasoning mechanism is subject to interfering factors. However, "whether the reasoning mechanisms are operating well or badly, we need not, and do not, assume that the individual to whom reasons are offered will exert any agency with respect to his or her beliefs" (131).

we do in relation to the ends we are working towards. In this way the agency in EA can be understood in purely epistemic terms although what the agent actually does cannot. I discuss each of these responses in turn.

With regards to EA requiring direct control of our beliefs, it will be helpful to contrast the two conceptions of EA under consideration. Kornblith's characterization of EA requires that one have direct control over lower-order belief formation if we are to demythologize EA. In other words, one must be able to have direct agency over one's beliefs. My suggested understanding of EA is more modest. The agency in EA is not the ability to control the lower-level processes, or mechanisms, of belief acquisition. Rather, the agency in EA takes place between the higher-level processes we can control and lower-level processes where belief simply happens. This conception does not require direct control of over our doxastic states.

Kornblith's airport screening agent example captures the distinction between our two views and highlights where I think Kornblith's conception is misplaced. The metal detecting wand represents the lower-level mechanistic belief-forming process; the ability to manipulate the wand by directing it in various directions represents the higher-level process. Kornblith argues that without the mechanical wand, no screening is possible. Because we do not control the mechanisms of the wand itself, wherever the wand is, *that* is where the actual screening takes place. Analogously, Kornblith argues that the actual formation of belief is at the mechanistic level, where there is no agency. Thus, he concludes that EA is a faulty concept.

This analogy does not work, however. Consider the possibility of there being both good and bad screeners. Poor screening could result from some mechanical issue with the wand itself or from the agent not knowing or adjusting the wand properly to the right kind of external stimuli. EA is not like simply directing the screening wand in any old way; rather, it is like learning to use the wand, or learning to use it better - or perhaps even like making improvements to the wand once one understands how it functions. We can allow that the end result takes place at the level of the wand's mechanisms, but we can also maintain that the screener has exercised agency over the screening process, and is at least partly responsible for successful screenings. Likewise for the agency one exercises over one's beliefs. The believer does not believe at will; the believer can, however, exercise agency over the belief-forming process. The crucial factor here is that what one does at the higher level can influence what occurs at the lower. This influence appropriately captures EA and shows how Kornblith's wand analogy doesn't actually support his conclusion.

EA, as I have developed it, then, is not subject to Kornblith's criticisms. His characterization of EA is a version of direct doxastic voluntarism, an epistemically unpopular view rarely endorsed by sympathizers of epistemic normativity or the possibility of doxastic improvement. So, while we are both addressing EA, the differences between our portrayals of this idea call into question whether we are indeed referring to the same concept. I suggest we are not, and that Kornblith is rejecting a notion that very few actually support.²⁹ To maintain that EA is mythological, then, Kornblith has to either (a) deny that each of the above agent-driven higher-order processes have any effect on our beliefs whatsoever or (b) provide us with reasons to accept his portrayal of EA. Kornblith agrees that if one believes p but acquires evidence for $\sim p$ after deliberation, or reflection, or some intentional action, this added evidence will affect one's doxastic state, even if the effect happens mechanistically. So option (a) is out. What about option (b)? Are there reasons that the agency in EA must be at the lower level? Kornblith does not give any reasons here. He simply asks: If the agency is not at the lower levels, then where? But as suggested above, this assumes that EA is some form of direct voluntarism about belief. So option (b) is out also.

It could be contended that this response does not do justice to the dialectic at hand. Kornblith accepts that there is agency occurring at the higher levels. Despite this type of agency, however, what one is doing when one is supposedly exercising EA can be accurately described without any distinctly epistemic terms. Believing, e.g., is not something that one does; belief is a phenomenon that happens independently of anything that an agent does *qua* agent. So there is nothing *epistemic* that is subject to the agency I suggest EA is based on. What we have, then, is simply a quibble over the term, or worse, an outright avoidance of the challenge that Kornblith presents to EA.³⁰

I do not deny Kornblith's challenge. I do think he is wrong to think that we are not agents with regards to what we believe and how we come to believe because there is a distinction between what one does as agent *qua* agent and how one's beliefs simply come about mechanistically. Kornblith's issue is that we

²⁹ There are numerous examples that could support this conclusion. Feldman, "The Ethics of Belief" and "Epistemic Deontologism," clearly rejects the suggestion that we have the ability to directly control our beliefs, while also accepting a normative component to what we believe. Nolfi also holds that there are prescriptive 'epistemic oughts' over which we have doxastic control, but not of the direct type espoused in DDV (Nolfi, "Why is Epistemic Evaluation Prescriptive"). Ryan likewise holds that we can be held epistemically and morally responsible for our beliefs while not having direct voluntary control over our beliefs (Ryan, "Moral Evidentialism").

³⁰ Thank you to the anonymous referees who highlighted these points.

needn't employ any distinctly epistemic terms when describing the agency involved in EA as I define it. This issue, however, leaves open another way in which we can think of agency, which does cite explicitly epistemic content: the goals and motivations we have when exercising EA.³¹

One of the central characteristics of EA is that one can exercise it when attempting to achieve a predominantly epistemic goal. Thus, it is not that we need to have the agency at the belief level; rather, it is that we have the agency with regards to achieving some epistemic ends. This proposal is obviously somewhat modest, but it does allow for other features that are evident in our own circumstances – like the ability to willfully improve in some epistemic ways. The practical versus epistemic agency distinction is certainly blurred here; indeed, as admitted above, EA is simply the exercise of our normal agency and thus there needn't be a distinction between the two. This conclusion does not matter. The agency exercised in EA reveals that we can be active with regards to what we believe. So while the activity is not the direct willing of a belief, the activity does affect how it is that we come to form and maintain some beliefs. And, while in some sense the dispute is merely terminological, it is nonetheless important as it reflects on deeper issues at work in the background. Taking Kornblith's challenge seriously, we can admit that there is no agency within the strictly epistemic domain, but this admission is not sufficient to deny EA. The agential focus is on achieving one's epistemic ends, like being able to form more accurate, fine-grained, sophisticated beliefs in general or in specific areas. If we can in some ways at some times achieve these goals – or even strive to achieve these goals – we are agents with regards to the results. If those goals are predominantly epistemic, then we are epistemic agents, even if what we do as agents is not directly controlling the epistemic.

Building from the preceding response, we can shift our focus to the issue of agential control and mechanistic bodily processes. Central to Kornblith's rejection of EA is that we have no control over these types of bodily processes. If we cannot directly control f , then we are not agents with regard to f . This assumption is not obviously warranted, however. There are marked differences between functions that we have some control over and functions that we have no control over. Contrast belief-formation with something like hair growth. Hair growth is mechanistic. We neither control the rate at which our hair grows nor the volume

³¹ Ahlstrom-Vij, whose rejection of EA we discuss below, recognizes this point as well: "...there is one type of account of [EA] that remains largely unaffected by Kornblith's critique, namely a type of account that takes [EA] to encompass the full range of things that we do in pursuit of epistemic goals..." (Ahlstrom-Vij, "Why We Cannot Rely," 277).

of hair we naturally have.³² There is a difference between the type of mechanistic processes over which we have no control, like hair growth, and those processes over which we have some control, like belief formation. Hair growth is not something we have any control over; how we form beliefs is something we have some control over.

Consider an analog to the type of indirect control we have over a mechanistic process like belief formation: heart rate. Whether or not one's heart beats and the rate at which it beats is not directly under one's control. That is, one cannot merely will one's heart to beat at, e.g., 72bpm. One can, however, undertake certain actions that we recognize will alter one's heart rate. If one sprints for one minute, one's heart rate will be significantly higher than if one were at rest. Furthermore, and more closely related to the type of control we have over our beliefs, one can affect one's overall heart rate indirectly via actions one can directly control. If I jog with the intention to improve my resting heart rate, and successfully do so over time, then I indirectly control the rate at which my heart will beat. I do not directly will my heart to beat at a certain rate; rather, I affect it more generally.

The simple appeal to mechanistic processes does not undermine the potential for intentionally affecting, albeit indirectly, that process in a goal-oriented way. We do not have the type of indirect control over hair growth that we do in bodily functions like our heart rates – and, as I am suggesting, in belief formation. There is a distinct difference between those processes that we have no control over whatsoever and those that we can indirectly control. Kornblith does not account for this difference; his emphasis is simply on the mechanistic aspects of the processes. There is a difference, however. And it provides further evidence that Kornblith's dismissal of EA is too hasty. That there is a mechanistic component to belief formation that we cannot directly control is not sufficient for the conclusion that we do not have some control over the beliefs we form. Neither is such lack of direct influence sufficient to show that we are not agents with regards to the ends we attempt to reach.

3.2 Ahlstrom-Vij's Rejection of Epistemic Improvement

Ahlstrom-Vij³³ provides a more direct challenge to EA. In contrast to Kornblith, Ahlstrom-Vij accepts that EA is more than simply controlling the lower-order processes of belief formation, but rather can be understood as the ways in which

³² There are, of course, surgical or other external ways to get more hair. But there are not steps that *I* can take in hair growth that are of my volition.

³³ Ahlstrom-Vij, "Why We Cannot Rely."

we attempt to achieve epistemic goals.³⁴ Given this characterization, which captures much of the presently defended conception of EA, Ahlstrom-Vij proposes that we should still reject this concept. His concern is that a central – perhaps the central – motivation in support of EA is that we want to and can partake in epistemic self-improvement. If indeed we are motivated to epistemically improve, “our main focus should be on the ways in which we fail *systematically*, rather than accidentally.”³⁵ That is, if we have systematic epistemic shortcomings, it is regarding those shortcomings that we should attempt to improve. Given this goal, however, we are faced with the problem that there are epistemic shortcomings that we are incapable of improving on, even if we want to. We are to conclude therefore that “we simply cannot rely on ourselves for epistemic improvement.”³⁶ If we cannot rely on ourselves for epistemic improvement, then we do not have EA.

The candidate example Ahlstrom-Vij employs to highlight our inability to epistemically improve involves cognitive biases, which are “systematic, and now well-established tendencies to form inaccurate beliefs.”³⁷ Consider, for example, the well-established cognitive biases involving tendencies to conflate personal traits when making comparative self-assessments. One may believe oneself to be more objective than 80% of a sample group; 60% of this group may assess themselves with the same ranking. At least a significant percentage of these assessments cannot be accurate as the ratio of people to percentiles does not jibe.³⁸ As this phenomenon represents a systematic epistemic failure on our part, we should be motivated to improve in this area. According to Ahlstrom-Vij, however, we cannot self-improve in these areas. We can run Ahlstrom-Vij’s argument thus:

1. If we have EA, then we can epistemically self-improve in areas where we systematically fail epistemically.
2. Cognitive biases are an area where we systematically fail epistemically.

³⁴ Ahlstrom-Vij assumes that “There is one and only one epistemic goal, and that is the dual goal of attaining true belief and avoiding false beliefs” (277). This is a general, all-encompassing, epistemic goal. The examples provided above concerning one’s goals to form more fine- or coarse-grained beliefs may be reducible to one horn of this goal. In any case, whether our respective portrayals of epistemic goals are consistent is inconsequential for the current discussion.

³⁵ Ahlstrom-Vij, “Why We Cannot Rely,” 278.

³⁶ Ahlstrom-Vij, “Why We Cannot Rely,” 293.

³⁷ Ahlstrom-Vij, “Why We Cannot Rely,” 278.

³⁸ Ahlstrom-Vij presents a number of psychological case studies where this type of phenomenon occurs.

3. We cannot epistemically self-improve our cognitive biases.

4. Therefore, we don't have EA.

Premise (1) establishes that we should be able to employ our EA to improve in areas of systematic epistemic failures. Premises (2) and (3), however, reveal an area of systematic failure where we cannot self improve epistemically. Thus, we are to conclude that we do not have EA.

Ahlstrom-Vij proposes three ways that we may attempt epistemic improvements in response to cognitive bias: self-correcting, self-binding, and external constraints. I focus on the first of these three proposals, as that is the focus of my response to the above argument. The self-correcting approach to cognitive bias suggests that we as epistemic agents can take steps to correct those systematic faults in reasoning that lead to inaccurate beliefs. According to Ahlstrom-Vij, there are two problems for the self-correction approach. First, there is a motivational problem resulting from our inability to recognize that we need to improve in this area in the first place – symptomatic of such biases is that we cannot accurately recognize when we are being biased in such a way.

Suppose, however, that we discover our cognitive biases. Could we not then take steps to correct them? Ahlstrom-Vij suggests that this approach won't work either, due to the *proper correction* problem. Even if we did somehow become privy to a tendency to conflate personal traits, self correcting via EA requires that we can do so successfully. Successful self-correction in turn requires that we satisfactorily deal with a number of challenges. The first challenge is that we recognize all and only the times that we are being biased. Another set of challenges are that the corrections made are all properly made. One has properly corrected if one neither over nor under corrects and one has made all and only the necessary corrections. Ahlstrom-Vij refers to numerous cases where at least one of these challenges is not met. While he accepts that the evidence he has presented against the self-correcting approach does not reject the possibility that one may be able to meet all of the above challenges, Ahlstrom-Vij highlights that *possible* does not mean *probable*. He concludes that

even if we assume that the relevant agents are at all motivated to engage in bias correction... there are substantial challenges they need to meet when it comes to doing so correctly.³⁹

What the above tells us is that there are good reasons to think that we cannot meet these challenges. So whether or not EA is marked by higher-order reasoning in conjunction with reasons-responsive lower-level bodily processes, we have

³⁹ Ahlstrom-Vij, "Why We Cannot Rely," 283.

reason to reject this concept as doing any substantial work. Its most appealing feature is undermined. That is, if we cannot epistemically self-improve in important ways – like in areas where failure is systematic rather than accidental – then there is reason to think we do not have EA.

Response to Ahlstrom-Vij

Ahlstrom-Vij's case against EA is compelling, but I think it falls short. The main problem arises in the first premise of his argument above: If we have EA, then we can undertake epistemic self-improvement in areas where we make systematic epistemic failures. If the failures are systematic, then it seems at the outset that we are doomed to fall short of fixing these failures. There are responses we might make to this premise, however. First, there might be areas where we cannot epistemically improve, but this inability in one area does not mean we cannot epistemically improve in other ways. Secondly, it's not actually clear that we cannot improve in the areas where cognitive bias manifests itself.

In support of the first response, in §2.3 above I gave a number of examples where epistemic improvement is possible and how we might go about improving in these ways. Admitting that we cannot improve when it comes to certain systematic biases we might have, does not mean admitting that we cannot epistemically improve *tout court*. It simply shows that there are limits to areas where we can improve. It is not controversial to accept that there are limitations on our abilities as epistemic agents. Indeed, there are limitations on our abilities as agents *qua* agents. Take the piano student who has learned all of the correct notes to Mozart's *Rondo Alla Turca*, but because of a systematic inability to keep time during high tempo songs cannot play it at a consistent speed. When it is pointed out that her timing is incorrect, she simply cannot improve. At slower tempos, however, she can keep time perfectly and can thus play Chopin's *Nocturnes*, e.g., as they were written. I suggest that it is an overgeneralization to propose that she is not an agent with respect to her piano-playing abilities because she systematically fails in one area that she is unable to improve on. Likewise for EA. There may be areas in which we are all systematically incapable of making epistemic self-improvements. All this tells us, however, is that we are not agents with respect to that part of our belief-forming abilities, not that we're not epistemic agents at all. So at most Ahlstrom-Vij highlights some limitation on EA.

A second response to Ahlstrom-Vij's first premise is to simply deny it by denying that the conditions needed for one to qualify as *improving* are too stringent. Recall that when taking steps to improve bias, Ahlstrom-Vij argues that one with EA would only correct for bias when one is actually being biased, would

not over or under correct, and would make all and only the necessary corrections. These conditions again seem reasonable in that they are the types of adjustments that would achieve the goal of forming the correct belief with the appropriate credence. We might ask, however, why do we need to have such fine-grained abilities? The standard seems too high; we do not place such stringent conditions on other less controversial exercises of agency. Consider the archer who misses the target 3' to the left. She adjusts her approach but overcorrects and misses the next shot 2' to the right. Although she over corrected, making more than the necessary corrections, most of us would admit that being a foot closer to the target is an improvement. Likewise if one discovers a personal epistemic shortcoming and overcorrects, but forms a belief that is closer to the truth than the previous one. This correction seems like an improvement. Or suppose one takes some but not all of the necessary steps to making a correct self-assessment in a case of cognitive bias. These steps also seem like improvements. Ahlstrom-Vij appears to have only provided us with idealizations for epistemic self-improvement, which might again simply highlight the limitations we have on becoming ideal epistemic agents.

One might counter that I am not being sensitive to the nuances of Ahlstrom-Vij's argument, most notably with what it takes to be an agent. If the archer, for example, takes aim at the target but has no notion of what the necessary steps are to hitting it aside from 'draw, aim, let go,' then the improvement between shots is merely a lucky accident. The necessary steps to hitting the target were not involved in her action plan and thus she was not an agent in respect to improving her shot – the improvement was mere luck. Similarly, the improvement in belief was lucky in the example concerning under or overcorrection or missing some of the necessary steps to eliminating cognitive bias. The apparent improvement made to achieve the goal of eliminating cognitive bias is also just luck.

Providing a full account of how lucky outcomes affect agency will take us too far astray from the point here. We can acknowledge, however, that one's agency is not undermined simply in virtue of a quasi-lucky outcome. As long as one is motivated to achieve one's goal, and proceeds with the intent to achieve that goal, one can still be an agent with regards to that undertaking, even if luck has some role in achieving the desired outcome. For example, one can intend to try to achieve one's goals despite there being uncertainty as to how to do so – or uncertainty as to whether or not one even can achieve the goal.⁴⁰ The archer can

⁴⁰ Alfred Mele, *Springs of Action* (Oxford: Oxford University Press, 1992), chapter 8, argues, e.g., that intentions to try to *A* can stand in for intending to *A*. One can be an agent with regard to achieving some ends, and thus achieve those ends intentionally, even if one's intentions were

intend to try to hit the target, even if she does not know all of the steps necessary for the successful completion of her goal to hit the target. Likewise, one can intentionally account for and improve cognitive biases in one's reasoning with the intention to try to form true and avoid false beliefs without knowing all of the necessary steps needed to do so, or without even being confident that one can.⁴¹ We have reason to accept, then, that even if an outcome is in some ways lucky, one can still intentionally improve in so far as they are motivated to improve and have taken steps – tried – to achieve that goal. I find it sufficiently plausible that intentionally doing anything denotes agency and so leave that undefended.

We have thus addressed reasons for rejecting Ahlstrom-Vij's rejection of EA. In the first case, if we grant that Ahlstrom-Vij is correct that we cannot improve in some cases of systematic epistemic failings, all we have been given is a limitation on where we can make epistemic self-improvements, not that we cannot make such improvements. In a second response, I propose that Ahlstrom-Vij's proper correction conditions are too stringent. If we held similar conditions for skill-based actions, we would undermine garden-variety agency. Failure to achieve a goal does not mean failure at making some improvements or that we are not agents with regard to those improvements. And even if we are limited in our ability to recognize the necessary steps to eliminate cognitive biases, or to what degree we need to adjust our self-assessments, that we can try to improve in these areas is sufficient to show that we have some agency here.

only to try to *A*. Michael E. Bratman, *Intentions, Plans, and Practical Reasons* (Cambridge Massachusetts: Harvard University Press, 1987), chapter 8, proposes a Single Phenomenon view of intentions and what one does intentionally: the intentional actions and acts done intentionally are both related to intentions, but that does not require that one intends what one does intentionally. In this case, one may intentionally improve by intending to try to improve although one might have constraints on their intending to improve. Hugh J. McCann, "Settled Objectives and Rational Constraints," *American Philosophical Quarterly* 26 (1991): 25-36, rejects Mele's and Bratman's distinction between intending to *A* and intending to try to *A*, arguing that intending to try to *A* is simply intending to *A*. Regardless of which view one accepts – whether or not one is intending to try to improve, or is intending to improve – if one does in fact improve, then one does so intentionally.

⁴¹ Donald Davidson, *Essays on Actions and Events* (Oxford: Oxford University Press, 1980), 92, provides a well-known example of intentionally doing something while also doubting whether or not one is or can be successful in what one is trying to achieve. In his example, Davidson is trying to make ten carbon copies by pushing his pen with some force while writing. He intentionally makes the copies, just in case he is successful in so doing; he is skeptical that he actually is making the copies while writing, however.

Conclusion

I have proposed that epistemic agency is best understood as the control we have over developing and refining our doxastic dispositions or belief-forming abilities, which include the propensity one has to form true or false or coarse-grained or fine-grained beliefs within different domains. In having this control we exercise epistemic agency in ways similar to indirect doxastic voluntarism – i.e. voluntarily undertaking a process so as to form a certain belief. Unlike indirect doxastic voluntarism, however, I have suggested that EA is not exercised so as to form a specific belief; EA merely affects the ways in which we can form beliefs more generally. This concept is motivated by the plausibility of epistemic normativity, skilled belief, and our ability to set and take steps towards achieving epistemic goals for doxastic self-improvement. There are good reasons, then, to accept EA as an actual phenomenon and a philosophically fruitful concept.

TRANSPARENCY AND REASONS FOR BELIEF

Benjamin WALD

ABSTRACT: Belief has a special connection to truth, a connection not shared by mental states like imagination. One way of capturing this connection is by the claim that belief aims at truth. Normativists argue that we should understand this claim as a normative claim about belief – beliefs ought to be true. A second important connection between belief and truth is revealed by the transparency of belief, i.e. the fact that, when I deliberate about what to believe, I can settle this deliberation only by appeal to considerations I take to show p to be true. It is natural to think that there is a connection between these two features of belief, that the fact that believing for non-evidential considerations would be irrational can help to explain why it is impossible, and Shah and Velleman make exactly this argument. However, as I shall argue, we cannot explain transparency on the basis of a normative requirement on belief. For this explanation to work non-evidential considerations would have to fail to be reasons for belief, and we would have to be able to explain why we are unable to form beliefs on the basis of non-evidential considerations by appealing to the fact that they fail to be reasons for belief. However, while it is plausible that non-evidential considerations are not in fact reasons for belief, the explanatory picture is the other way around. Such considerations only fail to be reasons for belief because we are unable to form beliefs on their basis.

KEYWORDS: transparency, aim of belief, truth, reasons, non-evidential considerations

Introduction

Belief has a special connection to truth, a connection not shared by other representational mental states such as imagination or supposition. We can see this connection in the fact that there is nothing out of the ordinary in saying “I am imagining that it is raining, but it isn’t raining,” but the Moore-paradoxical “I believe that it is raining, but it isn’t raining” is strikingly odd. One way that many philosophers have tried to cash the connection between belief and truth is in terms of the claim that, in some sense, belief aims at truth. However, there is substantial disagreement over the correct philosophical account of this claim. Beliefs, after all, are not themselves an agent who can have their own aims. According to normativists, the claim that beliefs aim at truth should be understood as a normative claim.¹ To say that belief aims at truth is to say that one

¹ Cf. Pascal Engel, “Doxastic Correctness,” *Proceedings of the Aristotelian Society* 87, (2013): 199-216, Pascal Engel, “In Defense of Normativism about the Aim of Belief,” in *The Aim of*

ought to hold a belief only if it is true, or that truth provides the standard of correctness for belief, for example.

Along with being the aim of belief, truth also has another interesting relation to belief. We can normally only form beliefs on the basis of considerations that we take to show the belief to be true. This is why, on proposing his wager as an argument for belief in God, Pascal goes on to recommend means by which one could bring oneself to actually form the belief that God exists. Accepting that belief in God is a good bet is not sufficient to bring about belief in God; you need to attend church, take communion, study the bible, and so on in the hopes of altering your evaluation of the truth of the belief in order to bring about this doxastic change. The situation is very different in cases where you become convinced that a consideration shows a claim to be true. In this case, we can form the belief directly, without adopting other means. This phenomenon has been called the transparency of belief. In some sense, the question of what to believe is transparent to the question of what is the case.² This explains why only evidential considerations, considerations that show the belief likely to be true in some way, can help us settle the question of what to believe.

It is natural to think that the aim of belief and the transparency of belief are related in some way. Both, after all, involve a special relationship between belief and truth. In particular, if we accept that the aim of belief should be understood in terms of a normative role for truth, then this looks like it should help us explain the transparency of belief. The fact that forming beliefs that aren't true is normatively forbidden may help us explain why it is impossible to do so, or at least to do so directly. Nishi Shah and David Velleman³ argue for just such an explanatory relation between the aim of belief and the transparency of belief. In fact, this explanatory link provides the main argument for their version of normativism about belief. For normativism to explain transparency, the agent

Belief, ed. Timothy Chan (New York: Oxford University Press, 2013), 32-63, Ralph Wedgwood, "The Aim of Belief," *Philosophical Perspectives* 16, (2002): 267-297, Nick Zangwill, "Directions of Fit and Normative Functionalism," *Philosophical studies* 91, 2 (1998): 173-203, Nick Zangwill, "The Normativity of the Mental," *Philosophical Explorations* 8, (2005): 1-19.

² This notion of transparency is related to but distinct from that described by Gareth Evans in *Varieties of Reference* (Oxford University Press, 1982). Evans focuses on the relation between truth and belief when we are forming beliefs, rather than in coming to know what we believe. I discuss how we should understand Velleman's notion of transparency in more detail in section 2.

³ See Nishi Shah, "How Truth Governs Belief," *The Philosophical Review* 112, 4, (2003): 447-482, Nishi Shah and J. David Velleman, "Doxastic Deliberation," *The Philosophical Review* 114, 4, (2005): 497-534.

must be aware of, or at least sensitive to, the normative requirements on belief. Otherwise, these normative requirements could not explain the psychological fact of transparency. This leads Shah and Velleman to argue that normativism is not (or not only) an independent normative truth, but part of the possession conditions for the concept of belief. In order to possess the concept of belief, Shah and Velleman argue, an agent must accept a normative claim, that beliefs are correct if and only if they are true. Thus, believers sophisticated enough to possess the concept of belief, and hence able to deliberate explicitly about what to believe, cannot help but be aware of the norm of truth, and this awareness can explain why beliefs formed through deliberation must be formed on the basis of considerations taken to be relevant to the truth of the belief.

However, tempting as it is, I do not think this explanatory strategy can ultimately be successful. We cannot explain transparency on the basis of a normative requirement on belief, not even if this norm is part of the possession conditions for the concept of belief. As I shall argue, for this explanation to work non-evidential considerations must fail to be reasons for belief. Furthermore, we must be able to explain why we are unable to form beliefs on the basis of non-evidential considerations by appealing to the fact that they fail to be reasons for belief. However, while it is plausible that non-evidential considerations are not in fact reasons for belief, the explanatory picture is the other way around. Such considerations only fail to be reasons for belief because we are unable to form beliefs on their basis. In other words, if we were able to form beliefs for non-evidential considerations, then such considerations would in fact count as perfectly valid reasons for belief. It is only our inability to actually believe for such reasons that prevents them from being reasons for us. And this shows that Shah and Velleman's strategy of explaining transparency in terms of a normative requirement fails.⁴ Furthermore, it provides strong reason to doubt that any similar explanatory strategy could succeed. This removes the main support for Shah and Velleman's theory of the aim of belief. But it also potentially has wider consequences. It remains quite plausible that the aim of belief and the transparency of belief have some kind of explanatory relation, and indeed that transparency is explained by the aim of belief. But if normativism cannot explain

⁴ For other criticisms of Shah and Velleman's position see Kathrin Glüer and Åsa Wikforss, "Against Belief Normativity," in *The Aim of Belief*, ed. Timothy Chan (New York: Oxford University Press, 2013), 80-99, Conor McHugh, "Normativism and Doxastic Deliberation," *Analytic Philosophy* 54, 4 (2013): 447-465, Andrei Buleandra, "Doxastic Transparency and Prescriptivity," *Dialectica* 63, 3 (2009): 325-332, Asbjørn Steglich-Petersen, "No Norm Needed: On the Aim of Belief," *The Philosophical Quarterly* 56, 225 (2006): 499-516, Asbjørn Steglich-Petersen, "Does Doxastic Transparency Support Evidentialism?" *Dialectica* 62, 4 (2008): 541-547.

Benjamin Wald

transparency, then this provides some reason to doubt that normativism is the correct account of the aim of belief.

1. Transparency

So what exactly is transparency? As Shah and Velleman put it,

The deliberative question whether to believe that p inevitably gives way to the factual question whether p , because the answer to the latter question will determine the answer to the former.⁵

In other words, when we deliberate about whether to believe p , we must settle our deliberation on the basis of exactly the same considerations that we would use to settle the question of whether p . Other considerations, although we might think about them and perhaps even wish we could form our belief on their basis, just do not settle the question of whether p , and so cannot settle the question of whether to believe that p either. We cannot, for instance, come to believe p because it would make us feel better, or because believing it would be good for our health, or because it would make our spouse happy.

This stands in stark contrast to how we deliberate about attitudes such as imagining or hypothesizing. We can decide to imagine that p , or hypothesize that p , for reasons that are utterly irrelevant to the truth of p . I can imagine that I have won an award just because imagining this would make me happy, but I cannot believe that I have won the award because the belief would make me happy. Thus, there is some special link here between belief and truth that shows up in our first personal deliberation about what to believe. It is important to note that this is not itself a normative claim – it is not that it is wrong to believe for pragmatic reasons, but that it is impossible to settle deliberation about what to believe by reference to anything other than evidential considerations, i.e. considerations we take to bear on the truth of the claim. There are actually two related claims being made here. The first is that one question, whether to believe that p , is transparent to a second question, whether p , when we deliberate. The second claim is that only what the agent takes to be evidence that p is true can be used by the agent to settle the question of whether p . However, I shall focus on the first claim, and take the second claim as given. The second claim will also gain some support from Shah and Velleman’s account of the nature of deliberation, discussed in section 3.⁶

⁵ Shah and Velleman, “Doxastic Deliberation,” 499.

⁶ The fact of transparency is not uncontroversial. Some philosophers think it is possible, and even sometimes rational, to give weight to non-evidential considerations in deliberating about what to believe (Cf. Conor McHugh, “The Illusion of Exclusivity,” *European Journal of*

Now, we should be clear about the strength of this transparency claim. Obviously, we might still be influenced in deciding whether to believe p by facts that have nothing to do with the truth of the belief. If I deliberate about whether to believe that my wife is cheating on me, I may be influenced by my deep desire not to believe this into discounting good evidence, in a way I would not have done had the question been about the faithfulness of someone else. What transparency rules out is that I could consciously decide not to believe that my wife is cheating on me on this basis. I can still be influenced by considerations that do not bear on the truth of the belief, but these must operate ‘behind the scenes,’ so to speak. The way in which these factors might influence my deliberation about whether or not to believe p is exactly the same way that they would operate in my deliberation about whether p , so we can retain the idea that the first question is transparent to the second.⁷ Furthermore, the claim that deliberation operates in this way is not merely armchair philosophical speculation. Psychological research on cases of so-called ‘motivated reasoning,’ where reasoners are incentivized to come to particular conclusions, suggests that the influence of practical incentives is indirect. Studies shows that, while people are in fact more likely to form a belief when they have been given practical incentives to form that belief, there is no conscious link between the non-evidential considerations and the formation of the belief. People spent longer looking at evidence that supported the belief they were incentivized to form, and spent longer searching their memory for instances that supported the desired belief,⁸ and the subjects were presumably unaware of this bias in their search for evidence. This provides empirical support for the claim that we can only form beliefs based on evidential considerations – when practical considerations affect our judgment, they do so by subconsciously affecting the way we look for or deliberate on evidential considerations, rather than by figuring explicitly in our deliberation. It is plausible that this is not just a contingent limitation on human believers, either. After all, a being that could form beliefs on the basis of non-evidential considerations could form beliefs on the basis of considerations they knew to be irrelevant to the truth of the belief. Thus they

Philosophy, forthcoming, DOI: 10.1111/ejop.12032). However, I shall assume for this paper that transparency is a real phenomenon.

⁷ We can also retain the claim that only considerations taken to be evidence for p can be used to settle the question of whether p , since the non-evidential factors are not being taken by the agent to settle the question of whether p , but instead unconsciously influencing the agent’s thinking about whether other considerations are good evidence for p .

⁸ Cf. Arie W. Kruglanski and Donna M. Webster, “Motivated Closing of the Mind: ‘Seizing’ and ‘Freezing,’” *Psychological Review* 103, 2 (1996): 263-283, Ziva Kunda, “The Case for Motivated Reasoning,” *Psychological Bulletin* 108, 3 (1990): 480-498.

Benjamin Wald

could form beliefs without regard to the truth of the belief, and such doxastic voluntarism is generally taken to be conceptually impossible,⁹ making transparency a conceptual truth about belief.¹⁰

2. Belief as a Normative Concept

Shah and Velleman argue that their version of normativism provides the best explanation of the phenomenon of doxastic transparency. They argue that the concept of belief has as part of its possession conditions the acceptance of a normative claim: namely, that beliefs are correct only if they are true.¹¹ Thus, to possess the concept of belief at all requires us to endorse a normative claim about when it is correct to hold a belief, so no believer can fail to be aware of this normative claim. The fact that believers necessarily endorse a norm for belief can be used, Shah and Velleman argue, to explain doxastic transparency.

Transparency, as Shah and Velleman understand it, only shows up when we deliberate about what to believe. As I said above, it is possible for belief to be influenced by non-evidential considerations, as long as these considerations operate behind the scenes. When beliefs are formed without deliberation, however, all of the influences on belief are similarly behind the scenes. It is only in deliberation that we explicitly consider what considerations count as reasons for forming a belief, and hence only here that there is a difference between the role of some considerations as reasons on which the belief is formed as opposed to mere causal influences in the formation of belief. Thus there is no distinction between the way that evidential and non-evidential considerations operate on beliefs formed without deliberation. However, when we deliberate explicitly about what to believe, only evidential considerations are relevant to settling the question. So what explains the fact that transparency shows up only when we deliberate? Well, because the deliberation is about what to *believe*, the agent

⁹ Jonathan Bennett, "Why is Belief Involuntary?" *Analysis* 50 (1990): 87-107, Bernard Williams, "Deciding to Believe," in *Language, Belief, and Metaphysics*, ed. Howard E. Kiefer and Milton K. Munitz (New York: SUNY Press, 1970), 95-111.

¹⁰ If transparency has a contingent psychological explanation then so much the worse for attempts to give a normative explanation of the phenomenon. However, I shall assume, in line with Shah and Velleman, that transparency is a conceptual truth.

¹¹ 'Correct' here is supposed to be a normative term, rather than a purely descriptive term. Several philosophers take the norms of belief to be given in terms of correctness. See, for example, Wedgwood, "The Aim of Belief," Engel, "In Defense of Normativism," and Alan Gibbard, "Truth and Correct Belief," *Philosophical Issues* 15 (2005): 338-351. However, nothing in the argument hinges on using correctness: for our purposes the result is the same if the norm of belief is given in different normative terms, such as what we ought to believe, instead.

necessarily applies the concept of belief in thinking about the outcome of deliberation. If Shah and Velleman are right about the possession conditions for this concept, then this entails that the agent endorses a norm that says that the belief which is the outcome of the deliberation will be correct if and only if it is true. Furthermore, for some mental process to count as deliberation, the agent must aim to reach the correct conclusion. A mental activity that was not aimed at reaching the correct outcome wouldn't count as deliberation at all – it would be idle imagining, perhaps, or even just a disconnected series of thoughts. So, in deliberating about what to believe, we are aiming to form a correct belief, due to the nature of deliberation, and we accept that only true beliefs are correct, due to the nature of belief. This, Shah and Velleman argue, shows that we are committed to forming the belief based only on factors we take to be relevant to its truth – just what doxastic transparency requires.¹²

However, I do not think this proposed account could truly explain transparency. To see why, consider what the strength of the proposed norm would have to be for it to explain transparency. Normally, the fact that a norm applies to something does not serve to constrain deliberation in the way that transparency does. Imagine that a friend asks me what I think of their haircut, and I judge that the new look is a colossal mistake, so I deliberate about what to say. Imagine further that I endorse a norm that forbids lying. Nonetheless, it seems that my deliberation could still include considerations such as the fact that telling the truth will hurt their feelings, and I might well end up choosing to act on this consideration, despite my acceptance of the norm against lying. The norm forbidding lying tells me that I should not say that p unless I think p is true, but this does not prevent me from taking into account or acting on considerations that have nothing to do with the truth of p . With transparency, on the other hand, these other considerations are prevented from having any influence. Thus, if transparency is explained by a norm, this norm must be of a special sort, unlike familiar norms such as the one forbidding lying.

We might try claiming that the norm of belief, unlike the norm against lying, is a *decisive* norm. While the norm against lying may provide some reason against lying, this reason still needs to be weighed against competing reasons to see if it is the strongest reason in this instance. However, perhaps the reason provided by the norm of belief is guaranteed to always be a decisive reason, outweighing any competing reasons. Thus, the agent has no need to consider

¹² Shah and Velleman also discuss in more detail their conception of the nature of deliberation and a mechanism for how we transition from deliberation to judgment and from there to belief, but the details of this account are not relevant to my criticism of it, so I omit them here.

other potential reasons for belief, since they can be sure that the reason provided by the norm of belief will always win out. However, this too falls short of accounting for transparency. Even if an agent knows that some consideration provides a decisive reason, it still seems possible for them to be swayed in their deliberation into acting for a different reason. Acting against what one takes to be a decisive norm is irrational, to be sure, but it is also a familiar phenomenon – if this weren't possible, then weakness of will would be much less prevalent. Consider again the norm against lying. Perhaps I have read a lot of Kant recently, and come to endorse the view that the norm against lying is a decisive norm, never outweighed by competing considerations. It still seems perfectly possible that, in a particular case, I might end up, through weakness of will, considering the harm to my friend's feelings, and acting on this basis. But in the case of doxastic deliberation, such weakness of will is not just irrational, but impossible. Not even a decisive norm seems to explain this impossibility.¹³

A final, and more promising, option is to hold that the norm of belief is a *silencing* norm. On this view, the norm of belief not only provides reasons that outweigh any competing reasons, it silences competing reasons, prevents them from having any rational weight at all. This entails that non-evidential considerations will fail to be reasons for belief. After all, the norm of belief is always in operation, and hence it will always silence non-evidential considerations. If they are always silenced, then non-evidential considerations will never have any weight in any deliberation about what to believe. But a consideration that never has any weight is thereby not a reason at all, so non-evidential considerations will not count as reasons for belief. Alternatively, we might think that, rather than the norm of belief silencing and hence eliminating competing reasons, there just never were any other reasons in the domain of belief in the first place. Perhaps the norm of belief provides the only reasons to be had when it comes to belief. These two explanations are structurally distinct, but the upshot is the same in either case – non-evidential considerations just do not count as reasons for belief.

If this were the correct interpretation of the strength of the proposed norm of belief, then Shah and Velleman would be arguing that possessing the concept of belief requires us to hold that the only things that count as reasons for belief at all are evidential considerations. This has better prospects of explaining why it is impossible, not just irrational, to form a belief on non-evidential considerations

¹³ This problem for Shah and Velleman's account has been previously noted by Steglich-Petersen, "No Norm Needed" and Sergio Tenenbaum, "Knowing the Good and Knowing What One is Doing," *Canadian Journal of Philosophy* 35 (2012): 91-117.

when deliberating. To believe on the basis of a non-evidential reason would, on this interpretation, involve deliberately forming a belief not just on the basis of a less pressing reason, but on the basis of something that is not even thought to be a reason at all. It seems plausible that this is not in fact possible. In the case of practical reason Joseph Raz¹⁴ points out that I cannot choose to have a coffee because I love Sophocles. If my love of Sophocles fails to in any way render my drinking coffee intelligible, i.e. fails to be a reason to drink coffee, then this consideration cannot be my reason for acting, and so cannot settle my deliberation about what to do. Similarly, if a candidate reason for believing *p* wouldn't render the formation of that belief at all intelligible, then it is plausible that it cannot be the agent's reason for believing *p*. The belief might be caused by the consideration, through some arational psychological process, but unless the consideration is seen as at least some reason for the belief, it couldn't count as the agent's conclusion in deliberation.

Furthermore, it seems plausible that this in fact how Shah thinks of the norm as functioning. For instance, he says that the effect of endorsing the norm of truth is that a strong disposition to block the influence of non-evidential types of influence is activated in cases of belief-formation that are governed by an agent's application of the concept of belief.¹⁵ This suppression of non-evidential considerations sounds more like a case of silencing such considerations than it does merely outweighing them. Similarly, Shah states that

belief's standard of correctness does determine what counts as a reason for belief from within the first-personal deliberative point of view.¹⁶

Shah here claims not only that the standard of correctness *provides* a reason, but also that it determines *what counts as a reason*. In other words, the claim of any other consideration to count as a reason at all depends on the norm of correctness, which suggests that it silences competing reasons.

This approach obviously requires that there in fact be no non-evidential reasons for belief. Furthermore, the argument requires that our inability to form beliefs for non-evidential reasons were explained by the non-existence of any such reasons. I shall argue in section four that, were we able to believe for non-evidential reasons, some of them would be perfectly good reasons for belief. This may seem to commit me to the unpopular view that there are in fact non-

¹⁴ Joseph Raz, "When We Are Ourselves: The Active and the Passive," in his *Engaging Reason: On the Theory of Value and Action* (New York: Oxford University Press, 2002), 5-21.

¹⁵ Shah, "How the Truth Governs Belief," 473.

¹⁶ Shah, "How the Truth Governs Belief," 472.

evidential reasons for belief. While this view is not obviously false,¹⁷ I am not committed to accepting this conclusion. Instead, we can hold that there are no non-evidential reasons for belief, but that the explanation of this fact is that it is impossible to form beliefs on the basis of non-evidential considerations. This, combined with a modest internalism about reasons for belief, entails that non-evidential considerations cannot be reasons for belief. However, this order of explanation will not help Shah and Velleman's argument, since they need the opposite order of explanation. In section five, then, I shall present a brief explanation of weak internalism about reasons for belief and show how, combined with transparency, it entails that there are no non-evidential reasons for belief, and also why this nonetheless is no help for Shah and Velleman. This will show how we can accept the arguments of section four without being committed to the existence of non-evidential reasons for belief.

3. Non-Evidential Reasons for Belief

For something to be a reason for belief implies that were we to form a belief on the basis of this consideration we would not be rationally criticisable, *ceteris paribus*,¹⁸ whereas we are rationally criticisable for forming beliefs on the basis of things that are not in fact normative reasons for belief. Thus, if we accept Shah and Velleman's claim that all believers must accept a norm that silences all non-evidential reasons for belief, then it should be rationally criticisable to form beliefs on the basis of these considerations. In this section, I shall argue that this is not true – if we were to form beliefs on the basis of some non-evidential considerations, this would not be rationally criticisable. Now, as I shall argue in section five, we might still hold that these considerations fail to be reasons for belief. In particular, they may fail to be reasons for belief precisely because we cannot form beliefs on the basis of such considerations. However, even so, it still remains true that were we able to form beliefs for these reasons, there would be nothing rationally criticisable about doing so.

¹⁷ For a defense of non-evidential reasons for belief, see Andrew Reisner, "The Possibility of Pragmatic Reasons for Belief and the Wrong Kind of Reasons Problem," *Philosophical Studies* 145, 2 (2009): 257-272.

¹⁸ The *ceteris paribus* clause here is important. We can be rationally criticisable for forming a belief on the basis of a genuine reason for belief if, for instance, there are stronger reasons against the belief, or a defeater for this reason is present. Still, there remains an important conceptual link between reasons for belief and rationality, which we can use to determine when a consideration counts as a reason for belief.

So why think that there are cases where it is not rationally criticisable to form a belief for non-evidential considerations? The basic argument is as follows. Opponents of non-evidential reasons for belief in general accept that non-evidential reasons can give us reasons to *bring it about* that we believe the proposition in question. For example, if an evil demon threatens to destroy the world unless you believe that the earth is flat, this provides you with a strong reason to take whatever means you can to bring it about that you believe this - read flat-earth arguments, try to convince yourself that there is a conspiracy against flat-earthers, get someone to hypnotize you, and so on. However, it seems very strange to say that it is rationally permissible to bring it about that you believe something, but were you able to bring yourself to believe the proposition directly, you would be rationally criticisable for doing so. Imagine someone who has the capacity to form beliefs on the basis of both evidential and non-evidential considerations, and who is deliberating about what to believe in the evil demon scenario. It seems highly implausible that he would be rationally criticisable for forming the belief that the world is flat on the basis of the demon's threat. Of course, we might object that forming the belief directly is not criticisable but impossible. I think this is exactly right, and perhaps, as I suggest in section five, we might think that as a matter of fact we would therefore have no reason to do as the demon commands, since it may be necessary for something to be a reason for belief that it is possible to form beliefs for this very reason. However, Shah and Velleman cannot appeal to this impossibility without rendering their position circular. Shah and Velleman need it to be impossible to form beliefs for non-evidential reasons because the agent would see it as violating the norm on belief that they must endorse to count as a believer; they cannot then explain the fact that it would violate the norm in terms of it being impossible to form the relevant belief. This is the core of my argument that it would not always be rationally criticisable to form a belief on the basis of a non-evidential consideration, and thus that there cannot be a general silencing norm forbidding forming beliefs for such reasons.

The counterfactual with which I frame my argument here may seem problematic. I argue that in some circumstance if we were able to form beliefs for non-evidential considerations then this would not be rationally criticisable. However, transparency is a conceptual truth, and hence necessary, and it states that we can only form beliefs for evidential considerations. So, the antecedent of this counterfactual is necessarily false, and thus according to the standard Lewisian semantics for counterfactuals the whole counterfactual is vacuously true. Similarly, the existence of an agent who could believe for non-evidential reasons

is similarly impossible, making the counterfactual framed in terms of such an agent also vacuous and uninformative.

I am not convinced that all counterfactuals with impossible antecedents are in fact vacuous. Consider, for instance, the claim that if Pythagoras' theorem were false, mathematicians wouldn't believe it. This seems non-vacuously true despite the necessary falsity of the antecedent.¹⁹ Still, I think we can make the same point without appealing to such counterpossible scenarios. Imagine instead an agent who is unsure of whether or not they can believe on the basis of non-evidential considerations. Perhaps they think they probably can't, but they aren't sure. When confronted by the evil demon considered above, this agent therefore tries to believe on the basis of the non-evidential considerations. If it is rationally criticisable to believe for non-evidential reasons, then it is rationally criticisable to try to believe for these reasons. But this agent does not seem rationally criticisable for making this attempt (although, depending on the scenario, they may be rationally criticisable for failing to recognize that it is impossible). On the other hand, imagine that this agent instead tries to believe that the earth is flat in order to annoy his philosophy teacher. In this case, the attempt does seem rationally criticisable, in a way that the attempt to satisfy the demon is not. While the goal is impossible in both cases, in the former the consideration speaks in favour of the belief, while in the latter it fails to do so. However, on Shah and Velleman's account, all such non-evidential considerations should equally be rationally criticisable to base beliefs on (or to try to do so), since all are equally ruled out as reasons by the aim of belief.

We might worry about the principle that, if it would be rationally criticisable to do something, then it is rationally criticisable to try to do it. But this principle is suggested by the plausible claim that agents are not rationally criticisable for failures caused by external factors over which they have no control. If I intend to visit Paris, I am rationally criticisable if I fail to buy a ticket or don't make plans to arrive at the airport on time, but I am not criticisable if the flight is cancelled due to a surprise storm. Once we subtract factors over which the agent has no control, however, trying one's hardest to do something and actually doing it are identical. To try one's hardest to do something is to do everything in one's

¹⁹ For further discussion of the view that counterpossibles can be non-vacuously true or false, and how to provide a semantics for them, see Jens Christian Bjerring, "On Counterpossibles," *Philosophical Studies* 168, 2 (2014): 327-353, Berit Brogaard and Joe Salerno, "Remarks on Counterpossibles," *Synthese* 190 (2013): 639-660, David Vander Laan, *Lewisian Themes: The Philosophy of David K. Lewis* (New York: Oxford University Press, 2004), Daniel Nolan, "Impossible Worlds: A Modest Approach," *Notre Dame Journal of Formal Logic* 38, 4 (1997): 535-572.

power to bring it about, and if external circumstances cooperate then one succeeds - there is no rationally relevant gap between the attempt and the success. This strongly suggests that the rational status of a successful action should be the same as a sufficiently determined attempt.

A defender of Shah and Velleman might object that I have not established the direction of explanation I am arguing for between reasons and belief. I have argued that it is our inability to believe on the basis of non-evidential reasons that explains their not being reasons, rather than the other way around. But, it could be objected, the counterfactuals I have made use of don't necessarily show this, even if true. They might instead show that if we were able to believe on the basis of these considerations that would be because they would, in that counterfactual situation, be reasons. Thus, a defender of Shah and Velleman's view could object that if I could respond to non-evidential considerations then they would be reasons, but not because my inability to respond explains their not being reasons. Instead, if I could respond to them, that would be because they were reasons.²⁰

However, it does not seem open to Shah and Velleman to claim that the considerations at issue are reasons in the counterfactual scenario described. Shah and Velleman are committed to the claim that it is a conceptual truth that only evidential considerations can be reasons for belief. Thus, in the scenario presented earlier, where an agent believes that the earth is flat in order to prevent the demon destroying the world, they would need to claim that the agent takes the demon's threat to be evidence that the world is flat. But it seems clear that the demon's threat is not evidence, and we can even add to the scenario that the agent doesn't take it to be evidence, and still generate the intuitive judgment that the agent's believing that the earth is flat on this basis would not be rationally criticisable. Thus, Shah and Velleman cannot offer as an explanation for the counterfactual the claim that the demon's threat is a reason in this scenario, since this would be to abandon the claim that reasons must, as a matter of conceptual necessity, be considerations the agent takes to be evidence.

Cases of 'motivated irrationality' might seem to provide examples of cases where it is in fact rationally permissible to do indirectly what it would be rationally criticisable to do directly. Parfit,²¹ for example, imagines a scenario in which a robber is trying to force you to open your safe so he can steal the gold, and he is willing to torture you or threaten your family in order to get you to comply. If you had a pill that would make you utterly irrational, Parfit argues, then the rational thing to do would be to take the pill. After all, if you were

²⁰ Thanks to an anonymous referee for pressing me on this point.

²¹ Derek Parfit, *Reasons and Persons* (New York: Oxford University Press, 1984).

utterly irrational, you would not respond rationally to the robber's threats, and so he would realize that there was no point in making these threats or carrying them out. Parfit describes the scenario playing out:

Reeling about the room, I say to the man: 'go ahead. I love my children. So please kill them.' The man tries to get the gold by torturing me. I cry out: 'this is agony. So please go on.' Given the state that I am in, the man is now powerless. He can do nothing that would force me to open the safe. Threats and torture cannot force concessions from someone is so irrational. The man can only flee, hoping to escape the police.²²

Cases of threats and deterrence may provide similar examples. If I can inculcate in myself a disposition to always carry out my threats, even when doing so is irrational,²³ this may be beneficial. Those I threaten, aware of my irrational disposition, might then accede to my demands, and thus I never have to actually carry through on my threats, so I end up benefitting.²⁴ In these situations it seems perfectly rational to make oneself irrational. However, performing the irrational actions directly would still be rationally criticisable. I would be rationally criticisable to directly act on my terrible threat, even if it is rational to bring about my disposition to carry out threats. This seems to provide a counter-example to the above argument, by suggesting cases in which it is rationally permissible to bring about what would be irrational to do directly.

However, these cases are importantly different from the scenario we began with. In the cases of motivated irrationality, what is rational to bring about is the *disposition* to perform irrational acts. We are not seeking to indirectly bring about an attitude or an action, but a disposition, and the benefit of the indirect actions is derived from the benefit of having this disposition. Performing the irrational actions directly would fail to realize this benefit. Carrying out a threat out of the blue fails to achieve the benefit of having the disposition to carry out threats, since the whole point of the disposition is deterrence. In the evil demon case we began with, the situation is different. What renders it rational to bring it about that I believe the earth is flat is the benefit of believing that the earth is flat; in particular, the fact that this belief will persuade the demon not to destroy the world. But forming the belief directly also achieves this very same result. It would be very odd if one and the same result could be achieved either directly or

²² Parfit, *Reasons and Persons*, 13.

²³ Due to the amount of harm carrying out my threat will lead to both for myself and for the threatened individual.

²⁴ Cf. Parfit, *Reasons and Persons*, David Gauthier, "Assure and Threaten," *Ethics* 104, 4 (1994): 690-721.

indirectly, and the benefit of the result makes the indirect method rational but not the direct method. It is not as if the direct method has terrible side effects that the indirect method lacks - the end result is identical, and if anything the indirect methods are more liable to produce undesirable side effects and cost extra time and effort.

But what if the evil demon threatens to destroy the world unless you believe something irrational? Wouldn't this show that sometimes one has a reason to indirectly bring something about, even though doing it directly would be irrational? After all, I have very good reason to bring it about that I believe something irrational, but by hypothesis I don't have any good reason to believe the irrational thing directly; if I did, it would be rational, and hence useless in my attempt to satisfy the demon. However, this is actually just another case where it is independently impossible to form a belief for a given reason; and hence, it has the wrong order of explanation to help Shah and Velleman. This is a bit easier to see in the practical case. Imagine that the demon has instead threatened to destroy the world unless I perform an irrational action. Imagine that, in order to comply with the demon, I hit myself in the head with a hammer as hard as I can. Is this irrational? It certainly would be normally. But in this case, if I am doing it because I believe this is the best way to prevent the demon destroying the world, then it in fact seems perfectly rational. Which, of course, defeats the point. So we have a conundrum. Almost anything I could do would be rendered rational by seeing it as a means to preventing the world being destroyed. And anything I shouldn't do even to save the world presumably still shouldn't be done. If I have a button that destroys the galaxy, then pressing this in order to prevent the world being destroyed would still be irrational, but only because pressing it is so much worse than the world being destroyed, so someone who does so is rationally (and morally) criticisable.²⁵ Thus, it seems that the only way to do something irrational is for that action not to be done in order to satisfy the demon, since this is a strong enough reason to render almost anything rational. But it seems plausible that the only way to do something irrational without it being done in order to save the world is to bring about the irrational action indirectly, perhaps by inculcating an irrational disposition in oneself and trying to forget the demon's threat altogether. However, this is not because there is anything wrong with the reasons I would be acting on if I acted directly. Preventing the demon from destroying the world is an

²⁵ The same goes for non-consequentialist reasons, although it is hard to think of examples of non-consequentialist reasons that are stronger than the reason in favour of saving the world. Still, if they exist, then presumably it would be wrong to violate this requirement even to save the world, so someone who does so is still rationally criticisable.

excellent reason to do something irrational, or would be if I could actually act on it. Sadly, precisely because it is such a strong reason, it is impossible to act for this reason. It will render the action I am trying to perform rational after all, defeating the purpose. Thus, here too, the order of explanation is wrong for this example to help Shah and Velleman. It is not that I lack sufficient reason to act irrationally, and this explains my inability to so act. Instead, only my inability to act for this reason prevents it from being an excellent reason. The same goes for the case of belief.

One could object at this point that there is at least some sense in which someone who intentionally forms a false belief in order to save the world is rationally criticisable. Even if we recognize the great practical benefits at stake, we might still say that they would be *epistemically* irrational to form a belief they took to be false.²⁶ The idea here is that epistemic reasons and practical reasons are not commensurable; they are two entirely separate standards of assessment. Epistemic reasons, on this view, are just those reasons that have to do with the truth or falsity of our beliefs, and we are epistemically irrational insofar as we fail to believe in accordance with these reasons. We can, of course, label a certain class of reasons as ‘epistemic reasons,’ and define corresponding notions of ‘epistemically rational’ and so on to accompany it. But this fails to address the main issue. The question of what to believe is a deliberative question facing agents. Recall our imagined agent who can deliberatively form beliefs on the basis of either evidential or non-evidential considerations. Such an agent would need to determine what they should believe. To tell them that there is one answer to what they should epistemically believe, and a different answer to what they should practically believe would be unhelpful- they would still be left with the unanswered question of what they should believe *simpliciter*. Imagine such an agent who is confronted by the evil demon who will destroy the world unless she believes the earth is flat. She knows the practical reasons favour believing that the world is flat, and the epistemic reasons favour believing that it is not flat, but she remains unsure what to believe. Is there really no further fact of the matter about what she should believe? This seems highly implausible.²⁷ Of course, we might

²⁶ It's not obvious that this solution will help Shah and Velleman. However, perhaps they could provide some further argument to explain why in deliberation we need to form beliefs not just for a reason, but for an epistemic reason. As I shall show, this move does not seem promising.

²⁷ This same point is frequently made about practical rationality. Some theorists about practical reason claim that there is no such thing as what ought to be done *simpliciter*, but only what ought to be done according to morality, what ought to be done according to self-interest, and so on (Cf. Philippa Foot, “Morality as a System of Hypothetical Imperatives,” *The Philosophical Review* 81, 3 (1972): 305-316, David Copp, “The Ring of Gyges: Overridingness and the Unity of

object that a believer who is capable of settling deliberation for either evidential or practical reasons is incoherent. The mental states formed by such a being just would not count as beliefs, we might suspect. But if so, this just shows that the true explanation of transparency lies with the explanation of why positing such a believer is incoherent, rather than with normativism about the aim of belief.

4. Internalism about Epistemic Reasons

The above argument suggests that there is nothing rationally criticisable about believing for non-evidential reasons. However, it does not necessarily follow that such considerations are in fact reasons for belief. Consider the following principle:

EPISTEMIC REASONS INTERNALISM: For some consideration p to be a reason to believe q , it must be possible for an agent to believe q for this very reason.

To unpack this claim, let us introduce the idea of a motivating reason for belief, by analogy with concept of a motivating reason for action. The motivating reason for one of my beliefs is, roughly, the consideration in light of which I form the belief, and also what I would appeal to if my belief were challenged.²⁸ Note that the way that I have described it, a motivating reason for belief is not usually a psychological fact, but instead a fact, or putative fact, about the world. After all, I would not usually appeal to my own psychology if challenged to defend one of my beliefs – I would appeal to what I took to be evidence for the belief.²⁹ We can

Reason,” *Social Philosophy and Policy* 14, 1 (2007): 86-106). However, many other theorists object that there must be such a thing as what we ought to do simpliciter, and that this is shown by the fact that, even after being told what morality recommends and what self-interest recommends, it is coherent and indeed natural to persist in asking what one ought to do (Cf. Stephen Darwall, “Reasons, Motives, and the Demands of Morality: An Introduction,” in *Moral Discourse and Practice: Some Philosophical Approaches*, ed. Stephen Darwall, Allan Gibbard, and Peter Railton (New York: Oxford University Press, 1997), 305-312, Sharon Street, “Reply to Copp: Naturalism, Normativity, and the Varieties of Realism Worth Worrying About,” *Philosophical Issues* 18 (2008): 207-228).

²⁸ This is not meant to be a definition of an explanatory reason for belief: it is intended to fix our attention on the appropriate phenomenon. For further discussion of motivating reasons in practical reason, see Michael Smith, *The Moral Problem* (Oxford: Blackwell Publishing Ltd. 1994), Jonathan Dancy, *Practical Reality* (Oxford: Clarendon Press, 2000), Kieran Setiya, *Reasons without Rationalism* (Princeton: Princeton University Press, 2010). For discussion of this same distinction applied to belief, see Kieran Setiya, “Epistemic Agency: Some Doubts,” *Philosophical Issues* 23 (2013): 179-198, Pamela Hieronymi, “The Wrong Kind of Reason,” *The Journal of Philosophy* 102, 9 (2005): 437- 457.

²⁹ Sometimes psychological facts may be motivating reasons for belief, as when I take the fact that I keep having sad thoughts as evidence that I am depressed, but this will not be the usual case.

contrast this with the idea of a normative reason for belief, which is the kind of consideration that actually counts in favour, normatively, of the belief. With this contrast in place, the suggestion under consideration is that for some fact to be a normative reason for belief, it must be capable of being a motivating reason for belief. This principle would be the theoretical analogue of a fairly weak form of internalism about practical reasons. Some philosophers hold that for something to be a normative reason for an agent that agent must be able to become motivated to act on the reason given their existing desires and psychology.³⁰ Our proposed principle is much weaker, since as I will argue the notion of possibility at issue is weaker than the kind of psychological possibility appealed to in these more strongly internalist arguments.

Epistemic reasons internalism may at first seem implausible. Consider someone who, perhaps due to very effective brainwashing in their youth, is unable to believe in the theory of evolution, and therefore unable to believe in the theory for the reason that it is supported by the best scientific evidence. This doesn't seem to show that this evidence thereby provides such a person with no reason to believe in the theory of evolution. The wealth of evidence for the theory still gives them very strong reason to believe it, even if they are unable to respond rationally to this evidence. However, I suspect that this is an issue of finding the correct notion of possibility. It may be psychologically impossible for the brainwashed individual to believe in the theory of evolution but this just shows that we should make use of a weaker form of possibility. The most plausible candidate is conceptual possibility. It must be at least conceptually possible for an agent to take some consideration as a motivating reason for belief for that consideration to be a normative reason for belief for that agent.

Furthermore, if we were right in claiming that transparency is a conceptual claim, then this will establish that non-evidential considerations cannot be reasons for belief. Of course, we will need some explanation of why transparency is a conceptual truth that does not, like Shah and Velleman's argument, rely on normativism. We might, for instance, argue that it is a conceptual truth about beliefs that they are mental states formed in response to evidential considerations. We can form representational mental states on the basis of considerations we do not take to bear on the truth of the content of such states, but these states will thereby fail to count as beliefs. They might instead be suppositions or

³⁰ Cf. Bernard Williams, "Internal and External Reasons," in his *Moral Luck*, (Cambridge: Cambridge University Press, 1981), 101-113, David Velleman, "The Possibility of Practical Reason," *Ethics* 106 (1996): 694-726.

imaginings. Pamela Hieronymi,³¹ for example, presents a non-normative argument that, if successful, would also establish that it is a conceptual truth that beliefs must be held for evidential reasons, by identifying the belief that *P* with the agent's answer to the question of whether *p*. Non-evidential considerations may make me wish that I could answer this question one way rather than another, but I can only actually settle the question, and thus form the belief, on the basis of considerations I take to be relevant to whether *p*, i.e. evidence. Shah and Velleman's argument relies on the idea that "deliberation is reasoning aimed at issuing in some result in accordance with norms for results of that kind"³². They then go on to argue that it is the norm governed aspect of deliberation that explains transparency - any agent who possesses the concept of belief must accept that truth is the only norm for beliefs. But we can use the characterization of deliberation given by Shah and Velleman to explain transparency even if we reject normativism. Deliberation not only aims to accord with norms for the result produced, it also aims at actually producing the result. Thus, deliberation about what to believe aims to produce belief. But if belief is, as a conceptual matter, something that must be formed on the basis of evidence, then deliberation about what to believe will be restricted to evidential considerations, because taking account of any other kind of consideration could not actually produce belief. If true, this account would show that there are no non-evidential reasons for belief, but this would be explained by transparency (together with the conceptual truths about deliberation and belief), and so could not be used to explain transparency without circularity. Thus, the truth of evidentialism on its own is not enough to save Shah and Velleman's argument. They need it to be the case that we cannot form beliefs for non-evidential reasons *because* there are no such reasons. However, epistemic reasons internalism need not establish this direction of explanation. Even if true, it might instead establish that there are no non-evidential reasons for belief because we cannot believe based on them.

Conclusion

Shah and Velleman's view, then, fails to adequately explain transparency. Non-evidential considerations would be reasons for belief, if we were able to form beliefs on the basis of such considerations. Thus, we cannot explain our inability to form such beliefs as a result of their not being reasons for belief - to do so would be circular. And this same argument suggests that the prospects for any

³¹ Hieronymi, "The Wrong Kinds of Reasons."

³² Nishi Shah, "A New Argument for Evidentialism," *The Philosophical Quarterly* 56, 225 (2006): 481-498.

Benjamin Wald

normative theory of the aim of belief being used as an explanation for transparency are dubious. This is, of course, a particular problem for Shah and Velleman's view. The purported ability of their theory to explain transparency served as the major argument for the view, so if this explanation fails the view is left largely unmotivated. However, I think this argument has implications for other views about the aim of belief. The original thought, that transparency has something to do with the aim of belief, remains highly compelling. Why, in deliberation, must our answer to the question of what to believe be resolved by our answer to the question of what is the case? Well, it seems plausible that it is because we are trying to form a belief in deliberating, and belief aims at truth. Absent an account of the aim of belief, this explanation is merely a sketch, but it seems to be on the right track. If, as I have suggested, we cannot explain transparency in terms of a norm of belief, then we will have to reject the suggestion that normativists can explain transparency by appeal to the aim of belief. Now, perhaps there is an explanation of transparency that has nothing to do with the aim of belief. I have certainly not said anything to rule out this possibility. But a theory that could account both for transparency and for the aim of belief seems like it would have a distinct advantage, and the inability of normativist understandings of the aim of belief to provide such a unified account is a mark against it.³³

³³ I would like to give special thanks to Sergio Tenenbaum, whose extensive comments on multiple drafts of this paper were invaluable. I would also like to thank Phil Clark, Elena Derksen, Mark Fortney, Jennifer Nagel, Luke Roelofs, Andrew Sepielli, audiences at the CPA annual congress and the University of Toronto grad forum, and two anonymous referees for very helpful comments on the paper.

BELIEVING AND ACTING: VOLUNTARY CONTROL AND THE PRAGMATIC THEORY OF BELIEF

Brian HEDDEN

ABSTRACT: I argue that an attractive theory about the metaphysics of belief – the pragmatic, interpretationist theory endorsed by Stalnaker, Lewis, and Dennett, among others – implies that agents have a novel form of voluntary control over their beliefs. According to the pragmatic picture, what it is to have a given belief is in part for that belief to be part of an optimal rationalization of your actions. Since you have voluntary control over your actions, and what actions you perform in part determines what beliefs you count as having, this theory entails that you have some voluntary control over your beliefs. However, the pragmatic picture doesn't entail that you can believe something as a result of intention to believe it. Nevertheless, I argue that the limited sort of voluntary control implied by the pragmatic picture may be of use in vindicating the deontological conception of epistemic justification.

KEYWORDS: rationality, pragmatism, doxastic voluntarism, epistemic deontology

Introduction

Do agents have voluntary control over what they believe? Philosophers almost universally hold that the answer is 'no.'¹ Instead, these philosophers maintain, belief formation is a passive affair, something that happens to the agent, as it were, when her evidence comes to support (or, in the bad case, when she merely takes her evidence to support) the proposition in question.

This negative conclusion certainly has the support of intuition. If you take some proposition (say, that it will rain tomorrow) and consider whether you are able to bring yourself to believe it, you will almost certainly conclude that you are not. But as I will argue, a mild and peculiar form of voluntary control over our beliefs follows from an attractive theory of the nature of belief, according to which what it is to have a belief is explained partly in terms of its link with action. According to such a pragmatic picture of belief, whether an agent has a given belief is determined in part by how she acts. In particular, whether an agent has a

¹ There are subtleties about direct vs. indirect, and immediate vs. long-term voluntary control, which will crop up later. But let's start out simple.

given belief is determined in part by how well attributing to her that belief would rationalize and explain her actions. Since an agent has voluntary control over her actions, and hence has control over one of the key factors that determine what she counts as believing, she thereby also has some voluntary control over what she believes. This, in a nutshell, is the argument from the pragmatic picture to the conclusion that agents have at least some voluntary control over what they believe.

While the pragmatic picture entails that we have some voluntary control over our beliefs, the form this voluntary control takes is peculiar. One might think that having voluntary control over what one believes involves being able (in some cases) to come to believe a proposition as a result of executing an intention to come to believe it. But, as I will argue, the pragmatic picture of belief will likely not have this consequence. On the view I sketch here, there are cases where if you act one way, you will count as having one set of beliefs, while if you act another way, you will count as having some other set of beliefs, and moreover it is under your voluntary control whether to act in the one way or the other. However, there are no cases where you can come to believe a proposition as a result of executing an intention or decision to believe that proposition. For in cases where you form the intention to believe p and subsequently act as if p is true, those subsequent actions will be better rationalized and explained by interpreting you not as believing p , but rather as merely desiring to believe p , or desiring to appear to believe p , or something of the sort. (For this reason, the resulting view may not merit the name ‘doxastic voluntarism,’ if this is taken to require voluntary control over beliefs via belief-forming intentions.)

Nevertheless, I will suggest that the more indirect kind of voluntary control over beliefs that follows from the pragmatic picture may be sufficient to defend what Alston² has called ‘the deontological conception of epistemic justification,’ on which beliefs are aptly evaluated using deontic concepts like *ought*, *obligation*, *permission*, and the like.

1. The Pragmatic Picture of Belief

The pragmatic theory of the nature of belief holds that what it is to have a given belief is for attributing that belief to you to be part of an optimal explanation of your behavior. This sort of pragmatic picture has been defended by Lewis,³

² See William Alston, “The Deontological Conception of Epistemic Justification,” *Philosophical Perspectives* 2 (1988): 257-99.

³ See David Lewis, “Radical Interpretation,” *Synthese* 27 (1974): 331-44.

Stalnaker,⁴ and Dennett,⁵ among others.⁶ Here is Stalnaker sketching the approach:

Here is one impressionistic picture of human activities which involve mental representation – call it the pragmatic picture. Rational creatures are essentially agents. Representational mental states should be understood primarily in terms of the role that they play in the characterization and explanation of action... And, according to this picture, our conceptions of belief and of attitudes pro and con are conceptions of states which explain why a rational agent does what he does. Some representational mental states – for example, idle wishes, passive hopes, and theoretical belief – may be connected only very indirectly with action, but all must be explained, according to the pragmatic picture, in terms of their connections with the explanation of rational action.⁷

One way to get a grip on the pragmatic theory of belief (and mental states more generally) is to imagine an ideal interpreter who knows everything about your behavior and environment and uses this knowledge to come up with a theory of your mental life. This ideal interpreter aims to attribute to you those mental states which together constitute the best explanation of your behavior. Whether some attribution of mental states to you is a good explanation of your behavior depends in large part on whether it is a good rationalization of your behavior, that is, whether it makes your actions come out rational. Crucially, this interpreter is not to be thought of as making hypotheses about what mental states you *really* have, hypotheses which could turn out to be false. Rather, what mental states you have at a time *just are* whatever mental states this ideal interpreter would attribute to you at that time. (Multiple competing attributions of mental states will typically be compatible with your acting as you do, but these may not all be equally good explanations and rationalizations of your actions. But if multiple competing attributions of mental states are tied with respect to how well they explain and rationalize your behavior, then it is indeterminate what mental states you have.)⁸

⁴ See Robert Stalnaker, *Inquiry* (Cambridge, MA: MIT Press, 1987).

⁵ See Daniel Dennett, *The Intentional Stance* (Cambridge, MA: MIT Press, 1987).

⁶ The pragmatic picture is sometimes referred to as ‘functionalism,’ but the term ‘functionalism’ has also been applied to views that do not qualify as versions of the pragmatic picture (e.g. views which simply espouse a computational theory of the mind). Still, we can say that the pragmatic picture is one version of functionalism about the mind.

⁷ Stalnaker, *Inquiry*, 4.

⁸ For discussion, see Alan Hájek and Michael Smithson, “Rationality and Indeterminate Probabilities,” *Synthese* 187 (2012): 33-48. They also highlight the parallel between Lewis’ theory of mental states and his theory of laws and chances.

As noted, the pragmatic theory of belief gives a central role to *principles of charity* which favor a given attribution of beliefs to an agent to the extent that such an attribution makes the agent come out rational. On sensible versions of the pragmatic theory of belief, making the agent's *actions* come out looking rational in light of her beliefs is just one part of making the agent come out rational as a whole (see Christensen⁹ and Meacham and Weisberg¹⁰ for related discussion). We also want the agent's beliefs to come out rational in light of her evidence,¹¹ for instance. And we want her beliefs to connect in sensible ways with her emotions, for instance whether she would experience surprise or disappointment if she were to learn that the proposition is false.¹² This is important, since there will typically be many different alternative sets of mental states that would recommend the actions the agent in fact takes (for instance, there are many credence-utility function pairs that assign highest expected utility to actions the agent performs), but these will not all be on a par with respect to how well they fit with the agent's evidence and emotions, for instance. It's also important to note that these different interpretational factors will sometimes conflict. For instance, it may be that the beliefs that would be most rational in light of the agent's evidence would not recommend the actions the agent in fact takes, or *vice versa*. Any particular version of the pragmatic theory will have to assign weights to the different factors, saying for instance how important it is that the beliefs attributed to the agent come out as rational in light of her evidence,¹³ how important it is that the beliefs

⁹ See David Christensen, "Preference-Based Arguments for Probabilism," *Philosophy of Science* 68 (2001): 356-76.

¹⁰ See Christopher Meacham and Jonathan Weisberg, "Representation Theorems and the Foundations of Decision Theory," *Australasian Journal of Philosophy* 89 (2011): 641-63.

¹¹ Note that insofar as an agent's evidence is determined by her mental states, a direct pragmatic picture of mental states in general (not just belief) will entail that what counts as her evidence is determined by the same explanatory considerations that determine what she counts as believing.

¹² The addition of these other constraints is part of what separates the sophisticated versions of the pragmatic picture advocated by Lewis, Stalnaker, and Dennett from implausibly crude versions. In the case of doxastic attitudes, such a crude version might take having a given credence in a proposition to be just a matter of betting at certain odds on that proposition. In the case of conative attitudes, a good example might be revealed preference theory, where preferring A to B just is a matter of choosing, or being disposed to choose, A over B when offered a choice between them. These ham-fisted views connect attitudes only with actions, and not with other things like what evidence one has, or what emotions one displays.

¹³ Because these different factors will often conflict, so that interpreters will often be unable to have agents come out as perfectly rational, the pragmatic theory will also need some measure of how far different suboptimal cases diverge from the ideally rational case. For instance, it will

attributed to the agent recommend the actions she performs, etc. The beliefs that the agent actually has, then, are those that score best overall, give the weights assigned by the theory to these different factors.

Still, what is crucial for present purposes is simply that what an agent believes is determined in part by how she acts. Lewis¹⁴ expresses the point with his Rationalization Principle, where **Ao** expresses what propositional attitudes an agent has, expressed in our language, **P** is our ultimate data set, described in purely physical terms, and Karl is the agent to whom we are attributing beliefs and desires:

The *Rationalization Principle*¹⁵ constrains the relation between **Ao** and **P**: Karl should be represented as a rational agent; the beliefs and desires ascribed to him by **Ao** should be such as to provide good reasons for his behavior, as given in physical terms by **P**. Thus if it is in **P** that Karl's arm goes up at a certain time, **Ao** should ascribe beliefs and desires according to which it is a good thing for his arm to go up then. I would hope to spell this out in decision-theoretic terms, as follows. Take a suitable set of mutually exclusive and jointly exhaustive propositions about Karl's behavior at any given time; of these alternatives, the one that comes true according to **P** should be the one (or: one of the ones) with maximum expected utility according to the total system of beliefs and desires ascribed to Karl at that time by **Ao**.

require a view not just about what beliefs would be perfectly rational in light of the agent's evidence, but also how irrational alternative beliefs states would be given that evidence, and it will require a view not just about what action would be most rational in light of a given set of beliefs and desires, but also how irrational other actions would be in light of those beliefs and desires. Note also that different versions of the pragmatic theory will also result from different theories of rationality. For instance, pragmatic theorists who are also causal decision theorists will sometimes disagree with pragmatic theorists who are evidential decision theorists about which beliefs would be an optimal rationalization of the agents' actions, since they disagree about the nature of rational action. And similarly for different theories about which beliefs are supported by which bodies of evidence. In this way, the pragmatic theory is more of a general picture of belief, with particular theories resulting from different ways of filling in that picture.

¹⁴ Lewis, "Radical Interpretation," 337.

¹⁵ Lewis' Rationalization Principle is one of the principles that I have called 'principles of charity.' In Lewis' terminology, he distinguishes between the Rationalization and the Principle of Charity. The latter says that agents should (*ceteris paribus*) be interpreted as having beliefs and desires which are themselves rational, while the former says that agents should (*ceteris paribus*) be interpreted as having beliefs and desires which make their actions rational in light of those beliefs and desires. Both the Principle of Charity and the Rationalization Principle have us aim to attribute beliefs and desires to an agent so as to make the agent come out as rational as possible overall.

If the pragmatic theory of belief is true, then you have voluntary control over some of the factors – your actions – which determine what you believe. If believing that p is in part a matter of acting as if p (i.e. acting in a way that would satisfy your desires if p were true), then by acting as if p you can contribute to making it the case that you believe p . This means that there will be possible cases in which, by acting one way rather than another, it will be the case that you believe some given proposition rather than not.¹⁶

A slight complication: On some versions of the pragmatic theory, it is not the agent's actions themselves which are in the first instance to be rationalized by the beliefs (and desires) attributed to her, but rather the agent's *dispositions* to act, which may or may not be manifested on any given occasion. I take it, however, that such a version of the pragmatic theory would still yield a limited form of voluntary control over beliefs, for not only the actions you actually perform, but also your dispositions to act in certain ways, are to some extent under your voluntary control. Henceforth I will ignore this minor complication and consider only the simpler version of the pragmatic theory on which it is the agent's actual actions which are to be rationalized by the beliefs (and desires) attributed to the agent.

The pragmatic picture of belief is structurally similar to the best-system analysis of laws of nature, and the way in which (I claim) agents have some voluntary control over what they believe has a structural analogue in the so-called 'undermining futures' which can arise in the context of objective chances in a best-system analysis. Here is Lewis describing the best-system analysis of laws:

Take all deductive systems whose theorems are true. Some are simpler, better systematized than others. Some are stronger, more informative, than others. These virtues compete: an uninformative system can be very simple, an unsystematized compendium of miscellaneous information can be very informative. The best system is the one that strikes as good a balance as truth will allow between simplicity and strength. How good a balance that is will depend

¹⁶ Another view of the metaphysics of belief that allows for a limited form of voluntary control over one's beliefs is one on which believing p is a matter of being disposed to treat P as a premise in practical reasoning. See Richard Holton, "Intention as a Model for Belief," in *Rational and Social Agency: Essays on the Philosophy of Michael Bratman*, ed. Manuel Vargas and Gideon Yaffe (New York: Oxford University Press, 2014) and Jacob Ross and Mark Schroeder, "Belief, Credence, and Pragmatic Encroachment," *Philosophy and Phenomenological Research* 88 (2014): 259-88. Treating something as a premise in reasoning is a sort of action, albeit a mental one. On this picture, then, your beliefs are under your voluntary control to the extent that your dispositions to perform certain actions are under your voluntary control.

Believing and Acting: Voluntary Control and the Pragmatic Theory of Belief

on how kind nature is. A regularity is a law iff it is a theorem of the best system.¹⁷

Note the similarity between the pragmatic picture of belief and the best-system analysis of laws. The best-system analysis says that what it is for something to be a law is for it to play a role in the best explanation of a certain pattern (namely, the complete world history). The pragmatic picture says that what it is for someone to have a given belief is for attribution of that belief to play a role in the best explanation of a certain pattern (namely, that agent's behavior). (An important difference is that what the laws are will not change over time, at least on standard theories, whereas what your beliefs are will change.)

Now I want to point out the analogy between the form of voluntary control over beliefs that follows from the pragmatic picture and the undermining futures that are possible given the best-system analysis of laws. First, let's look at how to incorporate objective chances into the best-system analysis. Here again is Lewis:

Consider deductive systems that pertain not only to what happens in history, but also to what the chances are of various outcomes in various situations - for instance, the decay probabilities for atoms of various isotopes... As before, some systems will be simpler than others. Almost as before, some will be stronger than others: some will say either what will happen or what the chances will be when situations of a certain kind arise, whereas others will fall silent both about the outcomes and about the chances. And further, some will fit the actual course of history better than others. That is, the chance of that course of history will be higher according to some systems than according to others... The virtues of simplicity, strength, and fit trade off. The best system is the system that gets the best balance of all three. As before, the laws are those regularities that are theorems of the best system. But now some of the laws are probabilistic. So now we can analyse chance: the chances are what the probabilistic laws of the best system say they are.¹⁸

On standard versions of the best-system analysis of laws, facts about what happens in the future are among the facts that determine what the laws are right now. For the same reason, on standard versions of the best-system analysis of chance, facts about what happens in the future are among the facts that determine what the chances are right now. As Lewis writes, this will be true if

present chances supervene upon the whole of history, future as well as present and past; but not upon the past and present alone.¹⁹

¹⁷ See David Lewis, "Humean Supervenience Debugged," *Mind* 103 (1994): 478.

¹⁸ Lewis, "Humean Supervenience Debugged," 480.

¹⁹ Lewis, "Humean Supervenience Debugged," 482.

Brian Hedden

If this is correct, then given what the present chances actually are, there is a non-zero chance that some future will obtain such that, if it were to obtain, the present chances would be different from what they actually are. To take Lewis' example, the actual half-life of tritium is 12.26 years (i.e. the chance that a given tritium atom will decay within 12.26 years is 0.5). But there is also

some minute present chance that far more tritium atoms will exist in the future than have existed hitherto, and each one of them will decay in only a few minutes.²⁰

If this were to happen, then it would constitute a

chancemaking pattern on which the half-life of tritium would be very much less than the actual 12.26 years.²¹

Now, the possibility of such 'undermining futures' is of interest to philosophers of science primarily because it creates a conflict between the best-system analysis and plausible claims about how one ought to respond to evidence about objective chances, especially the Principal Principle of Lewis.²² But for present purposes, what is important is just the observation that on the best-system analysis of chance, facts about how things go right now and in the future can make a difference to what the present chances are. And by the same token, on a pragmatic picture of belief, facts about how you act right now and in the future (at least, the very near future) can make a difference to what your present beliefs are.

Suppose, for instance, that for each of a number of tritium atoms, you are offered a bunch of bets at various odds on whether the tritium atom will decay within 12.26 years. In particular, for each tritium atom, and for all n between 0 and 1 (inclusive) have the option of accepting or declining a bet which pays you $\$n$ if the tritium atom decays within 12.26 years and pays you $\$(n-1)$ otherwise. Note that having credence x that the given tritium atom will decay would license accepting all and only the bets where n is greater than or equal to x (in the sense that that bet would have non-negative expected value, given that credence).²³ If, then, you accept all and only the bets where n is greater than or equal to 0.5, this behavioral pattern is one that would be best explained and rationalized by your having 0.5 credence that a given tritium atom would decay within 12.26 years. By

²⁰ Lewis, "Humean Supervenience Debugged," 482.

²¹ Lewis, "Humean Supervenience Debugged," 482.

²² See David Lewis, "A Subjectivist's Guide to Objective Chance," in *Studies in Inductive Logic and Probability*, ed. Richard Jeffrey (Berkeley: University of California Press, 1980), 83-132.

²³ I'm assuming that you only care about money, and that you don't have decreasing marginal utility for money. The story gets more complicated once we drop these simplifying assumptions, but the basic point is the same.

contrast, if you accept all and only the bets where n is greater than or equal to 0.25, this behavioral pattern is one that would be best explained by your having only 0.25 credence that a given tritium atom would decay within 12.26 years. Of course, there are other factors that, on the pragmatic picture, help determine what your beliefs (or credences) are, for instance what evidence you have encountered, how surprised you would feel if you learned that the given proposition is false, and the like. But given that how you act is among the factors that determine what you believe, it is difficult to resist the conclusion that there will be at least one possible case where which bets you accept at time t affects what credences you have at t regarding tritium atom decay (just imagine a case where the two alternative beliefs states are on a par with respect to these other factors, so that your actions can tip the balance one way or the other). Since how you bet is (presumably) under your voluntary control, this would also be a case where what you believe, indeed what you believe *right now*, about tritium atom decay, is to some extent under your voluntary control.

Importantly, this sort of case is unlike a case in which you cause your later self to have a given belief by taking some belief-inducing pill or by giving your later self new evidence (Feldman²⁴ notes that you can cause your later self to believe that the lights are on by flipping the light switch). This is a case where you can affect what you believe *right now* by how you act *right now*. By accepting all and only the bets that appear fair or favorable relative to credence n in a given proposition, you make it the case that you count as having that credence n right before you accepted the bets (here assuming that other interpretational factors, like whether that credence would be supported by the evidence you had then, do not conflict with this assignment of credences).

Note that we can get an analogue of undermining futures on the pragmatic picture. Suppose that in fact you accept all and only the bets where n is greater than or equal to 0.5, so that your actions are best explained and rationalized by your having 0.5 credence that a given tritium atom will decay within 12.26 years. And suppose that the other factors in the supervenience base for your beliefs/credences are neutral with respect to which credence-attribution they support (so, for instance, you have not been exposed to information that would rationalize one particular credence over another). So, assuming that the pragmatic picture is correct, you in fact have 0.5 credence that a given tritium atom will decay within 12.26 years. However, had you acted in some other way, say by accepting all and only the bets where n is greater than or equal to 0.25, then you

²⁴ See Richard Feldman, "The Ethics of Belief," *Philosophy and Phenomenological Research* 60 (2000): 667-95.

Brian Hedden

would have had some credence other than the 0.5 credence that you actually have. Therefore, just as the best-system analysis of chance allows for possible futures such that, if they obtained, the chances would be different from what they actually are, so the pragmatic picture allows for possible courses of action such that, if you had performed them, your beliefs would be different from what they actually are.

It is important to note that the sort of voluntary control over one's beliefs that I am claiming follows from the pragmatic picture of belief is very weak. This is precisely because how you act right now is only one factor among many that determine what you count as believing. Again, other factors include how you acted in the past and how you will act later in the future, as well as what information you have encountered and what emotions or other responses you are disposed to have under various conditions. And these other factors are not (or at least, not typically) right now under your voluntary control. In many cases, these other factors will point strongly in one direction, supporting some particular attribution of beliefs to you, such that even if you acted in a way that would not be rationalized by those beliefs, you would still count as having them. Your action would then count as a case of one-off irrationality.

For instance, if you have seen lots of polling that shows the Republicans leading in the midterms, then if you were to nonetheless bet (at even odds) on the Democrats, this would not be enough to make it the case that you believed that the Democrats would win. For the fact that you have strong evidence that the Republicans will win means that attributing to you the belief that the Democrats will win would require interpreting you in such a way that your beliefs turn out highly irrational (even though it would make your actions come out rational in light of your beliefs). Of course, we do sometimes interpret agents as having irrational beliefs; it's just that the principle of charity says we should try to avoid that result if possible (subject to other constraints). So, if you have strong evidence that the Republicans will win, then if you still bet on the Democrats, it would be natural to interpret you as believing the Republicans will win, with your betting actions counting as irrational, or perhaps as expressive acts showing your support for the Democrats, or an attempt to trick others into thinking you believed the Democrats would win, or something of that sort.

Only one of the components of the supervenience base of your beliefs, namely, how you act right now, is under your present voluntary control. It is for this reason that even on the pragmatic theory, cases where you can affect what

you believe right now by how you act right now will be rare. I only claim that they are possible.²⁵

Let me close this section with a few words about what the pragmatic picture says about voluntary control over mental states other than belief. The pragmatic picture of Lewis, Stalnaker, and Dennett is not merely a theory about belief, but also a theory about a range of other propositional attitudes,²⁶ in particular desires (or preferences) and intentions. For instance, whether an agent counts as having a certain set of desires (or preferences) depends on whether those desires would rationalize and explain her actions. (As with beliefs, though, it will also depend on whether those desires would themselves be rational, and on how those desires connect with her emotional dispositions, for instance whether she would

²⁵ Daniel Greco, in “The Impossibility of Skepticism,” *Philosophical Review* 121 (2012): 317-58, distinguishes between direct and indirect pragmatic picture of belief. The direct pragmatic picture is the one I have been considering. On the indirect pragmatic picture, the link between belief and action is (as the name suggests) more indirect. As he puts it, the indirect pragmatic picture has it that “a proposition *p* counts as believed by an agent if a representation with the content that *p* is produced by a belief-producing psychological mechanism of that agent,” where “for a system that produces representations to count as a *belief*-producing system, it must have the function of producing representations that play the action-guiding, rationalizing role that the advocate of the direct pragmatic picture thinks beliefs must play” (337). Whether some form of voluntary control follows from the indirect pragmatic picture depends on what determines the function of a representation-producing system. If the function of such a system is fixed by the past evolutionary, selective history of that system (yielding something like the biosemantic picture of belief defended in Ruth Millikan, “Biosemantics,” *Journal of Philosophy* 86 (1989): 281-97), then no form of voluntary control will follow, for the function of the relevant representation-producing system is fixed by facts about the past which are not under her present control. On the other hand, if the function of a representation-producing system is determined by, say, species-wide regularities (e.g., whether in typical members of the species, the representations it produces play a belief-like role in guiding and rationalizing action), then a very weak form of voluntary control will follow. This is because an agent’s present actions play a role in determining whether the representations produced by a given representation-producing system play a belief-like role in guiding action. Hence, her actions play a role in determining whether that representation-producing system counts as a *belief*-producing system. Then, in marginal cases, she might be able to tip the balance and make one of the representation-producing systems in her head count as a belief-producing system, and thereby make it the case that she counts as believing whatever representations that system produced. These cases will be exceedingly rare, and so the form of voluntary control that follows from this version of the indirect pragmatic picture will be even weaker than that which follows from the direct pragmatic picture. This is unsurprising, since the link between belief and action is weaker and less direct on the indirect pragmatic picture than on the direct pragmatic picture.

²⁶ Recall that in the quote from Stalnaker above, he mentions ‘idle wishes, passive hopes, and theoretical belief’ in addition to ordinary beliefs.

experience a feeling of disappointment if those desires were frustrated.) And whether an agent counts as having some intention depends on whether that intention would rationalize her actions. Because these other mental states are also explained in part by their connection with the explanation of rational action, an agent will in some cases be able to determine which of those mental states she has by how she acts, and hence will have some voluntary control over those mental states.

By contrast, it is more questionable whether the pragmatic picture applies to mental states with a strong phenomenal component. It is questionable, for instance, whether it applies to states of having certain qualia, like there seeming to be a red thing in front of me.²⁷ Plausibly, these states are characterized by their subjective feel, not (or at least not to any great extent) by their connection to rational action. A more difficult intermediate case is that of certain emotions like fear, anger, and surprise, which involve a phenomenal component but are not mere phenomenal states. Griffiths²⁸ argues that certain ‘core’ emotions, such as fear, surprise, anger, disgust, sadness, and joy, are automatic responses to environmental stimuli characterized by their associated physiological changes and their evolutionary phylogeny²⁹ the same emotions, associated with similar physiological changes, in other animals such as dogs), rather than by their connection with rational action. (However, Griffiths doesn’t think that this story applies to certain higher cognitive emotions like jealousy and guilt, for instance.) If that story about the core emotions is on the right track, then the pragmatic picture will not entail that they are in any sense under our voluntary control, because the pragmatic picture doesn’t apply to them; what core emotions one counts as having doesn’t depend on whether attribution of those emotional states would explain and rationalize one’s actions, but rather on facts about physiology and evolutionary history. Of course, there is much more to be said about phenomenal and emotional states, and their connection to the pragmatic picture. My point is simply that the pragmatic picture doesn’t entail that all of our mental states are in any sense under our voluntary control, since the pragmatic picture likely doesn’t apply to all mental states, but only those characterized by their connection with rational action.

²⁷ See Jaegwon Kim, *Physicalism, or Something near Enough* (Princeton: Princeton University Press, 2007).

²⁸ See Paul Griffiths, *What Emotions Really Are* (Chicago: University of Chicago Press, 1997).

²⁹ See Charles Darwin, *The Expression of the Emotions in Man and Animals* (Originally published 1872. Reprinted by University of Chicago Press, 1965).

2. Voluntary Control and the Deontological Conception

I have argued that on a pragmatic picture of belief, your beliefs are partly under your voluntary control, since part of the supervenience base for your beliefs, namely your actions, are under your voluntary control. This is an interesting result, in part because the form this voluntary control takes is peculiar and unexpected. But does it have any broader relevance? In this section, I discuss how the pragmatic picture of belief bears on one topic of central importance in discussions of voluntary control over beliefs, namely the deontological conception of epistemic justification.

According to the deontological conception of epistemic justification, justification has to do with requirements, prohibitions, and permissions, so that (for instance), a belief is justified just in case it is permitted by the relevant epistemic norms. Thus, beliefs can appropriately be evaluated using ordinary deontic concepts such as *obligation*, *permission*, *ought*, and the like, and it's appropriate to say that an agent is obligated to hold a certain belief, given her evidence, or that a certain belief is permissible, or that an agent ought not have the beliefs she has, and so forth. Alston³⁰ criticizes this conception of epistemic justification by first arguing that beliefs must be under some sort of voluntary control in order for deontic concepts such as *requirement*, *prohibition*, and *permission* to apply to them (by appeal to the principle that *ought* implies *can*), and then arguing that we lack the requisite voluntary control over our beliefs.

Of course, whether Alston's argument is sound depends on what sort of voluntary control is needed in order for the relevant deontic concepts to be applicable to beliefs. And it is by no means clear and uncontroversial what sort of voluntary control is at issue. (A closely related issue is what it would take to believe 'at will,' and Peels³¹ identifies eight different conceptions of believing at will that have been discussed in the literature.)

An initial thought, and one to which Alston seems sympathetic, gives pride of place to intentions (or, perhaps, choices or decisions).³² On this view, having the relevant sort of voluntary control over whether you ϕ is a matter of being able to ϕ as a result of executing an intention to ϕ (and, presumably, being able to not- ϕ as a result of executing an intention to not- ϕ). So, having the relevant sort of

³⁰ Alston, "The Deontological Conception."

³¹ See Rik Peels, "Believing at Will is Possible," *Australasian Journal of Philosophy* 88 (2014): 1-18.

³² Alston ("The Deontological Conception," 259) writes that according to the principle that *ought* implies *can*, "one can be obliged to do A only if one has an effective choice as to whether to do A."

voluntary control over whether you believe p is a matter of being able to believe p as a result of an intention to believe p , and being able not to believe p as a result of an intention not to believe p .³³

But in my earlier discussion, I said nothing about intentions to believe. For all that I have said, it might be that in some cases, by acting one way rather than another, you can make it the case that you have one set of beliefs rather than another, even though it is impossible for you to form an intention to believe p and have that intention cause you to come to believe p . After all, in the cases I have imagined, by acting one way rather than another, you make it the case that you *already* had the one set of beliefs rather than the other. So our ideal interpreter would not interpret you as having intended to have the one set of beliefs and *thereby* causing yourself to acquire those beliefs.

In this respect, there is another illuminating analogy between the sort of voluntary control that I think follows from the pragmatic picture and Lewis' best-system analysis of laws. Lewis is a compatibilist about free will. Lewis³⁴ imagines that determinism is true and that in fact he put his hand down on his desk and did not raise it. This action was predetermined but nonetheless free. It was predetermined since there is a true historical proposition h specifying the intrinsic state of the world long ago, and a true proposition I specifying the (deterministic) laws of nature, and h and I together entail that Lewis did not raise his hand. But Lewis was free in keeping his hand down. He was able to raise his hand; he just didn't exercise that ability.

Now, Lewis asks what would have been the case had he raised his hand. There are three possibilities. Either contradictions would have been true, or h would have been false, or I would have been false. The first is easily dismissed. Lewis also dismisses the second, noting that if he had raised his hand, the intrinsic state of the world long ago would not have been different. Only the third option remains. Had Lewis raised his hand, I would have been false. The laws of nature would have been different from what they actually are.

Now for the crucial part. Facing another incredulous stare, Lewis addresses the worry that this description of the case means that he is able to break the laws

³³ Kieran Setiya, in "Believing at Will," *Midwest Studies in Philosophy* 32 (2008): 36-52, endorses an intention-based conception of believing at will, on which being able to believe at will involves being able to believe as a result of an intention to so believe. Then, if one takes the deontological conception of epistemic justification to require agents to be able to believe at will, we get the result that it requires agents to be able to form beliefs as a result of intentions to form those beliefs.

³⁴ See David Lewis, "Are We Free to Break the Laws?" *Theoria* 47 (1981): 113-21.

of nature. Lewis replies that he is not committed to the strong thesis that he is able to break a law, but only to the weak thesis that he is able to do something such that, if he did it, something that a law in the actual world would have been broken. In the nearest world in which Lewis raises his hand, it is not the case that something which is a law of nature in that world is broken. Rather it is the case that something which is a law of nature in the actual world would be broken, and hence was never a law in the nearest Lewis-hand-raising world in the first place.

Something similar is going on in the belief case. Suppose, to use our earlier example, you actually accept all and only bets on tritium atom decay where n is greater than or equal to 0.5. Given our stipulation that other factors on which your beliefs supervene are neutral with respect to which beliefs or credences you count as having, this means that you actually have 0.5 credence that a given tritium atom will decay within 12.26 years. But you could have bet otherwise, accepting all and only the bets where n is greater than or equal to 0.25. Had you done so, you would have counted as having 0.25 credence that a given tritium atom will decay within 12.26 years. This is a case where your ability to affect your beliefs is somewhat akin to Lewis' ability to break the laws. In the nearest world in which you accept all and only bets where n is greater than or equal to 0.25, it is not true in that world that you caused your present credence in a tritium atom's decaying within 12.26 years to go from 0.5 to 0.25 as a result of a decision you made to come to have 0.25 credence in that proposition. Rather, in that world, you did something such that, given that you did it, you count as having already had the 0.25 credence in the first place.

So all that follows from the pragmatic picture of belief is that there are cases where at t you are able to do something such that, if you do it, you will count as having one set of beliefs at t , and you are able to something else such that, if you do that other thing, you will count as having some other set of beliefs at t . It does not follow from the pragmatic picture of belief that there are cases where you can be interpreted as having decided or intended to believe some proposition p and thereby coming to believe p . The pragmatic picture of belief does not straightforwardly rule out the latter possibility; it just doesn't by itself entail that it is possible. Whether coming to believe p as a result of deciding to do so is possible on the pragmatic theory depends on the specifics on the version of the pragmatic theory in question. In particular, it will depend on the requirements of rationality that our imagined ideal interpreter seeks to interpret agents as satisfying, and on the different weights assigned to these different requirements in cases where satisfying one requirement entails violating another.

However, let me briefly explain why I suspect that on most plausible versions of the pragmatic theory, it will come out impossible to believe some proposition p as a result of an intention or decision to believe p . Why should this be impossible? Couldn't you intend at t to believe p at a slightly later time $t+\epsilon$, know that ϕ -ing at $t+\epsilon$ would be best explained and rationalized by your believing p at $t+\epsilon$, and then ϕ at $t+\epsilon$, thereby bringing it about that you believe p at $t+\epsilon$? Arguably not. For holding fixed that you intended at t to believe p at $t+\epsilon$, it is likely that your subsequent ϕ -ing at $t+\epsilon$ would be best explained and rationalized not by your *believing* p at $t+\epsilon$, but rather by your merely *desiring* to believe p at $t+\epsilon$, or perhaps desiring to appear to others as if you believed p . The interpreter would attribute to you a desire to believe p (or something along those lines) while refraining from attributing to you the belief itself. (Alternatively, our interpreter might attribute to you the belief that p at $t+\epsilon$ but then not interpret you as having intended at t to come to believe p . After all, as noted earlier, most defenders of the pragmatic picture will hold that for many if not all contentful mental states, including not just beliefs but also intentions, whether you have that mental state is determined by how well attributing to you that mental state would rationalize and explain your behavior.)³⁵

Suppose that this is right and the pragmatic picture does in fact rule out the possibility of intending or deciding to believe some proposition and thereby

³⁵ There are some worries about this explanation of why it should be impossible to form beliefs as a result of an intention to do so. Presumably, the idea behind thinking that an interpreter would prefer interpreting you as desiring to believe p at $t+\epsilon$ rather than as actually believing p at that time is that, given the rest of your situation, the belief that p wouldn't be rational. But why should that be the case? Perhaps it's because such a belief would have to be based on merely pragmatic, non-evidential reasons. But even if such beliefs are irrational, it might be that in the case under consideration, you have evidential, non-pragmatic reasons to decide to try to get yourself to come to believe p . For instance, you might recognize that your evidence supports p but find yourself unable to believe p just as a result of your ordinary passive belief-formation processes. You might recognize that your evidence suggests that climate change is caused by human activities but find that your anti-intellectual upbringing is keeping you from responding in the normal way to this evidence and believing in man-made climate change. In such a case, intending to get yourself to come to believe in man-made climate change would involve responding to evidential rather than pragmatic reasons. Second, even if it irrational to believe on the basis of pragmatic reasons, it is not clear that this irrationality is so egregious as to outweigh any explanatory benefits that might be gained by interpreting you as having done so. The pragmatic theory allows for the possibility of interpreting you as having irrational beliefs or performing irrational actions in other contexts, so why not this one? I leave these questions open at present.

coming to believe it. This yields a form of doxastic voluntarism on which your ability to determine what you believe parallels the ability that Lewis thinks you have to break the laws. There will be cases where by doing one thing at t , you will count as having one set of beliefs at t , and by doing some other thing, you will count as having some other set of beliefs, even though it is impossible for you to have some set of beliefs as a result of an intention to do so.

However, it is possible to resist the assumption that the deontological conception of epistemic justification requires the sort of voluntary control in which you can form beliefs on the basis of belief-forming intentions. In fact, Alston himself considers a different kind of voluntary control that may suffice for defending the deontological conception. He calls it 'indirect voluntary control.'³⁶ Here is how he puts it:

We can be held responsible for a state of affairs that results from our actions even if we did not produce that state of affairs intentionally, provided it is the case that something we did (didn't do) and should have not done (done) was a necessary condition (in the circumstances) of the realisation of that state of affairs, i.e., provided that state of affairs would not have obtained had we not done (done) something we should not have done (done)... This suggests that even if propositional attitudes are not under our effective voluntary control, we might still be held responsible for them, provided we could and should have prevented them; provided there is something we could and should have done such that if we had done it we would not have had the attitude in question.³⁷

In this passage, Alston is suggesting that the deontological conception of epistemic justification may only require that we have voluntary control over (some of) our beliefs in the following sense: one has voluntary control over whether believes p iff one has voluntary control over whether one does something (intentionally) which is a necessary condition (in the circumstances) of believing p .

But *this* kind of voluntary control over beliefs is precisely the kind that follows from the pragmatic picture of belief (or so I have argued). Take a case where one will count as believing p if and only if one ϕ 's. Then, one has voluntary control over whether one believes p , provided one also has the ability to (intentionally) ϕ and the ability to (intentionally) not- ϕ . For one has voluntary control over whether one does something (namely ϕ) which is a necessary condition (in the circumstances) of believing p .

³⁶ Alston, "The Deontological Conception," 277.

³⁷ Alston, "The Deontological Conception," 278.

Admittedly, Alston didn't have in mind the peculiar sort of voluntary control over beliefs that follows from the pragmatic picture when he wrote his article. He was conceiving of our indirect voluntary control over beliefs as stemming from our ability to take actions (he mentions training oneself to be more critical of gossip as an example) that are within one's power to do and which causally affect one's later belief formation. By contrast, the indirect voluntary control over beliefs resulting from the pragmatic picture is one on which we are able to take actions which *constitutively* make a difference to one's beliefs. Still, insofar as indirect voluntary control over beliefs, as characterized by Alston in the quote above, is sufficient to defend the deontological conception of epistemic justification, the pragmatic picture of belief can provide refuge to epistemic deontologists.

Now, even if one insists that this indirect sort of voluntary control doesn't merit the name and isn't voluntary control at all, I think that the pragmatic picture can still help vindicate our doxastic responsibility. (We might then be thought of as having a form of 'doxastic freedom' while lacking voluntary control over beliefs.³⁸) Suppose that in fact your evidence supports believing p , and that if p is true, then ϕ -ing is the action that would best satisfy your desires.³⁹ Let us also suppose that you in fact don't ϕ ; instead you ψ . And attributing to you a lack of belief in p (either a belief in not- p or suspension of judgment) is part of the best explanation and rationalization of your ψ -ing. And suppose further that, were you to have ϕ -ed, then attributing to you the belief that p would have been part of the best explanation of rationalization of your ϕ -ing (after all, this belief would itself be rational, i.e. supported by your evidence, and would rationalize your actions). Then, we can say that you ought to have believed that p , and that you were free to have done so. For had you acted in the way that your evidence suggested was best (i.e. had you ϕ -ed), then you would have counted as believing that p , and you had the freedom to ϕ .⁴⁰

³⁸ See Sharon Ryan, "Doxastic Compatibilism and the Ethics of Belief," *Philosophical Studies* 114 (2003): 47-79, Conor McHugh, "Exercising Doxastic Freedom," *Philosophy and Phenomenological Research* 88 (2014): 1-37, and Matthias Steup, "Doxastic Voluntarism and Epistemic Deontology," *Acta Analytica* 15 (2000): 25-56.

³⁹ I'm holding fixed what your desires are for the sake of this example, though according to the pragmatic picture, what your desires are also depends in part on how you act.

⁴⁰ Again, it won't always be the case that you would have counted as believing differently had you acted differently, for in some cases other facts having to do with your evidence and your emotions, for instance, will overdetermine what you count as believing. But most philosophers, including Alston ("The Deontological Conception," 262) have thought that you needn't *always* have freedom or voluntary control over your beliefs in order for deontic concepts to be

Conclusion

Lack of voluntary control over beliefs is largely taken for granted in contemporary philosophy. But many philosophers also espouse theories of the nature of belief that involve a tight conceptual link between belief and action. In particular, many hold that what it is for an agent to have a given belief is in part for her to act in ways that would satisfy her desires if that belief were true. More generally, they hold that what it is for you to have a given set of mental states, including beliefs, is for attribution of those mental states to constitute (part of) an optimal explanation of your behavior. On such a pragmatic theory of (many) mental states, your actions are part of the supervenience base for your mental states, including your beliefs. But given that your actions are under your control, this means that some of the factors on which your beliefs supervene are under your control, and so there are possible cases in which you can affect what you count as believing by how you act. These cases are likely rare, and perhaps even non-actual, but they are possible nonetheless.

However, the sense in which your beliefs can in some cases be under your control differs from the sense in which some philosophers often think of voluntary control. The pragmatic theory of belief does not entail that there are possible cases in which you affect what you believe by way of a decision or intention to come to have that belief. Rather, the form of voluntary control over your beliefs entailed by the pragmatic theory of belief is like the form of voluntary control over the laws of nature entailed by Lewis' compatibilism. According to the latter, you are able to do something such that, were you to do it, the laws of nature would be different from what they would be if you didn't, but there is no world in which you break the laws of nature in that world. According to the former, you are able to do something such that, were you to do it, your beliefs would be different from what they would be if you didn't, but there may be no world in which you form a given belief as a result of a decision to do so.

Nevertheless, I have suggested that the sort of voluntary control that follows from the pragmatic picture (which might be thought of as a form of Alston's 'indirect voluntary control') may be enough to vindicate our doxastic responsibility and hence the deontological conception of epistemic justification.

applicable to beliefs; it is enough that you sometimes, or at least often, have such freedom or voluntary control.

NOTES ON THE CONTRIBUTORS

Patrick Bondy is with the Department of Philosophy at Trent University, on a Limited Term Appointment as Assistant Professor. Prior to that, he held a Postdoctoral Fellowship with the Sage School of Philosophy at Cornell University. His main research interests are in epistemology and argumentation theory. His recent and forthcoming publications include “Epistemic Value” (*Internet Encyclopedia of Philosophy*, forthcoming), “Counterfactuals and Epistemic Basing Relations” (*Pacific Philosophical Quarterly*, forthcoming), “Virtues, Evidence, and Ad Hominem Arguments” (*Informal Logic*, forthcoming), and “Epistemic Deontology and Strong Doxastic Voluntarism: A Defense” (*Dialogue*, 2015). Contact: patrbondy@gmail.com.

Brian Hedden received his PhD from the Massachusetts Institute of Technology in 2012 and is a Lecturer in Philosophy at the University of Sydney. He works mainly in epistemology and decision theory, with interests in philosophy of language, ethics, and philosophy of science. He recently published a book entitled *Reasons without Persons* (Oxford University Press, 2015) and is the author of numerous journal articles. Contact: brian.hedden@sydney.edu.au.

Dustin Olson is a PhD candidate at the University of Rochester. His current research defends an epistemic interpretation of reflective equilibrium and applies this interpretation to address challenges found in belief control, epistemic normativity, and peer disagreement. Additional research interests include the philosophy of action, the philosophy of time, social and political thought, and Bertrand Russell. Contact: dustin.olson@rochester.edu.

Andrew Reisner is a Senior Lecturer in Practical Philosophy at Uppsala University. Prior to that he was an Associate Professor of Philosophy at McGill University. He has published on a range of topics in theoretical and practical reason, metaethics, and value theory. He is co-editor, with Iwao Hirose, of *Weighing and Reasoning: Themes from the Work of John Broome* (Oxford University Press, 2015), and he is also co-editor, with Absjörn Steglich-Petersen, of *Reasons for Belief* (Cambridge University Press, 2011). Contact: andrew.reisner@filosofi.uu.se.

Sharon Ryan is Associate Professor of Philosophy at West Virginia University. Her research interests are in epistemology, metaphysics, ethics, and philosophy for children. A sample of her publications include: “Wisdom” (*Stanford Encyclopedia of Philosophy*, 2013), “Wisdom, Knowledge, and Rationality” (*Acta Analytica*, 2012), “Doxastic Compatibilism and the Ethics of Belief” (*Philosophical Studies*, 2003), “What is Wisdom?” (*Philosophical Studies*, 1999), “The Epistemic Virtues of Consistency” (*Synthese*, 1996), “Does Warrant Entail Truth?” (*Philosophy and Phenomenological Research*, 1996). She is also creator of THE QUESTION (thequestion.blogs.wvu.edu, 2004-2010), an online project designed to get everyone, including young children, excited about philosophy. Contact: sharon.ryan@mail.wvu.edu.

Joseph Van Weelden is a PhD candidate in Philosophy at McGill University. His research is primarily in value theory, specifically the theory of prudential value. His current project sets out to develop and defend a new version of pluralism about prudential value. His philosophical interests span much of contemporary ethics and metaethics. Secondary areas of interest include epistemology (especially testimony) and metaphysics (especially personal identity). Contact: joseph.vanweelden@mail.mcgill.ca.

Benjamin Wald is a PhD candidate in philosophy at the University of Toronto. His research interests include philosophy of action, ethics, epistemology, and practical and theoretical reasoning. His current research project is to defend the view that acting for a reason involves taking one’s motivating reason to be a normative reason, and showing how this view fits into an attractive picture of both the nature of belief and intention and the nature of practical and theoretical inference. Contact: Benjamin.wald@utoronto.ca.

LOGOS & EPISTEME: AIMS and SCOPE

Logos & Episteme is a quarterly open-access international journal of epistemology that appears in printed and on-line version at the end of March, June, September, and December. Its fundamental mission is to support and publish various current reflections and researches that aim at investigating, analyzing, interpreting or philosophically explaining the human knowledge in all its aspects, forms, types, dimensions or practices.

For this purpose, the journal publishes articles, reviews or discussion notes focused as well on problems concerning the general theory of knowledge, as on problems specific to the philosophy, methodology and ethics of science, philosophical logic, metaphilosophy, moral epistemology, epistemology of art, epistemology of religion, social or political epistemology, epistemology of communication. Studies in the history of science and of the philosophy of knowledge, or studies in the sociology of knowledge, cognitive psychology, and cognitive science are also welcome.

The journal promotes all methods, perspectives and traditions in the philosophical analysis of knowledge, from the normative to the naturalistic and experimental, and from the Anglo-American to the Continental or Eastern.

The journal accepts for publication texts in English, French and German, which satisfy the norms of clarity and rigour in exposition and argumentation.

NOTES TO CONTRIBUTORS

1. Accepted Papers

The journal accepts for publication articles, discussion notes and book reviews submitted exclusively to *Logos & Episteme* and not published, in whole or substantial part, elsewhere. The editors of *Logos & Episteme* reserve the right to refuse all future papers belonging to the authors who fail to observe this rule.

2. Submission Address

Please submit your manuscripts electronically at: logosandepisteme@yahoo.com. Authors will receive an e-mail confirming the submission. All subsequent correspondence with the authors will be carried via e-mail. When a paper is co-written, only one author should be identified as the corresponding author.

3. Paper Size

The articles should normally not exceed 12000 words in length, including footnotes and references. Articles exceeding 12000 words will be accepted only occasionally and upon a reasonable justification from their authors. The discussion notes must be no longer than 3000 words and the book reviews must not exceed 4000 words, including footnotes and references. The editors reserve the right to ask the authors to shorten their texts when necessary.

4. Manuscript Format

Manuscripts should be formatted in Rich Text Format file (*.rtf) or Microsoft Word document (*.docx) and must be double-spaced, including quotes and footnotes, in 12 point Times New Roman font. Where manuscripts contain special symbols, characters and diagrams, the authors are advised to also submit their paper in PDF format. Each page must be numbered and footnotes should be numbered consecutively in the main body of the text and appear at footer of page. For all references authors must use the Humanities style, as it is presented in *The Chicago Manual of Style*, 15th edition. Large quotations should be set off clearly, by indenting the left margin of the manuscript or by using a smaller font size. Double quotation marks should be used for direct quotations and single quotation marks should be used for quotations within quotations and for words or phrases used in a special sense.

5. Official Languages

The official languages of the journal are: English, French and German. Authors who submit papers not written in their native language are advised to have the article checked for style and grammar by a native speaker. Articles which are not linguistically acceptable may be rejected.

6. Abstract

All submitted articles must have a short abstract not exceeding 200 words in English and 3 to 6 keywords. The abstract must not contain any undefined abbreviations or unspecified references. Authors are asked to compile their manuscripts in the following order: title; abstract; keywords; main text; appendices (as appropriate); references.

7. Author's CV

A short CV including the author's affiliation and professional address must be sent in a separate file. All special acknowledgements on behalf of the authors must not appear in the submitted text and should be sent in the separate file. When the manuscript is accepted for publication in the journal, the special acknowledgement will be included in a footnote on the first page of the paper.

8. Review Process

The reason for these requests is that all articles, with the exception of articles from the invited contributors, will be subject to a strict blind-review process. Therefore the authors should avoid in their manuscripts any mention to their previous work or use an impersonal or neutral form when referring to it. The review process is intended to take no more than six months. Authors not receiving any answer during the mentioned period are kindly asked to get in contact with the editors. Processing of papers in languages other than English may take longer. The authors will be notified by the editors via e-mail about the acceptance or rejection of their papers. The editors reserve their right to ask the authors to revise their papers and the right to require reformatting of accepted manuscripts if they do not meet the norms of the journal.

9. Acceptance of the Papers

The editorial committee has the final decision on the acceptance of the papers. Papers accepted will be published, as far as possible, in the order in which they are

received and they will appear in the journal in the alphabetical order of their authors.

10. Responsibilities

Authors bear full responsibility for the contents of their own contributions. The opinions expressed in the texts published do not necessarily express the views of the editors. It is the responsibility of the author to obtain written permission for quotations from unpublished material, or for all quotations that exceed the limits provided in the copyright regulations. The papers containing racist and sexist opinions assumed by their authors will be rejected. The presence in texts of sexist or racist terms is accepted only if they occur in quotations or as examples.

11. Checking Proofs

Authors should retain a copy of their paper against which to check proofs. The final proofs will be sent to the corresponding author in PDF format. The author must send an answer within 3 days. Only minor corrections are accepted and should be sent in a separate file as an e-mail attachment.

12. Reviews

Authors who wish to have their books reviewed in the journal should send them at the following address:

Eugen Huzum
Institutul de Cercetări Economice și Sociale „Gh. Zane”
Academia Română, Filiala Iași
Str. Teodor Codrescu, Nr. 2, 700481
Iași, România.

The authors of the books are asked to give a valid e-mail address where they will be notified concerning the publishing of a review of their book in our journal. The editors do not guarantee that all the books sent will be reviewed in the journal. The books sent for reviews will not be returned.

13. Property & Royalties

Articles accepted for publication will become the property of Logos & Episteme and may not be reprinted or translated without the previous notification to the editors. No manuscripts will be returned to their authors. The journal does not pay

royalties. Authors of accepted manuscripts will receive free of charge one copy of the issue containing their papers.

14. Permissions

Authors have the right to use their papers in whole and in part for non-commercial purposes. They do not need to ask permission to re-publish their papers but they are kindly asked to inform the Editorial Board of their intention and to provide acknowledgement of the original publication in *Logos & Episteme*, including the title of the article, the journal name, volume, issue number, page number and year of publication. All articles are free for anybody to read and download. They can also be distributed, copied and transmitted on the web, but only for non-commercial purposes, and provided that the journal copyright is acknowledged.