

Strong Internalism, Doxastic Involuntarism, and the Costs of Compatibilism

Timothy Perrine

Synthese, DIO: 10.1007/s11229-018-1879-4

Abstract: Epistemic deontology maintains that our beliefs and degrees of belief are open to deontic evaluations—evaluations of what we ought to believe or may not believe. Some philosophers endorse strong internalist versions of epistemic deontology on which agents can always access what determines the deontic status of their beliefs and degrees of belief. This paper articulates a new challenge for strong internalist versions of epistemic deontology. Any version of epistemic deontology must face William Alston’s argument. Alston combined a broadly voluntarist conception of responsibility, on which ought implies can, with doxastic involuntarism, the position that our beliefs are not under our control. Together, those views imply that epistemic deontology is false. A promising response to Alston’s argument is to embrace a compatibilist account of control—specifically a reason-responsive version of compatibilism—and use it to criticize his doxastic involuntarism. I argue that while reason-responsive compatibilism about control does undermine Alston’s argument, it comes at a cost. Specifically, it is inconsistent with strong internalist versions of epistemic deontology. The surprising upshot is that so long as we retain a voluntarist conception of responsibility, we have reason for rejecting strong internalist versions of epistemic deontology.

Broadly construed, epistemic deontology is the position that our degrees of belief and our doxastic attitudes—our beliefs, disbeliefs, and withholdings—are subject to deontic appraisal.¹ That is, epistemic deontology is the position that it makes sense to evaluate (e.g.) our beliefs in terms of what we *ought* to believe, or are *permitted* to believe, or *may not* believe.² Epistemic deontology implies that our doxastic attitudes or degrees of belief are subject to epistemic obligations, even if it fails to specify what those obligations are.³

Internalism about epistemic deontology is a family of positions that provides an internalist account of epistemic obligations. One version of internalism is what I call “strong internalism.” It maintains that the only facts that determine the deontological status of an agent’s doxastic attitudes are facts that the agent can access. In this way, it urges that whatever goes on beyond the purview of the agent’s perspective is irrelevant to the deontic appraisals of the agent’s beliefs.

¹ I assume both beliefs and degrees of belief exist. For purposes here, I am agnostic as to whether one category reduces to the other. I will focus mostly on beliefs, but all the issues here could be formulated *mutantis mundandi* in terms of degrees of belief.

² These are not the only possible evaluations of beliefs. We can evaluate beliefs in terms of their value or goodness, epistemic virtuous or viciousness, coherence, origins in a reliable source, whether they have reasons or evidence that support them, etc. (And, of course, there might be various relations between these evaluations—a subject I will not discuss here.)

³ There is a category of ‘ought to be,’ e.g. “it ought to be that hardworking students get jobs.” Some might propose securing deontic evaluations for agents’ doxastic attitudes by appeal to this category (see, e.g., Chrisman (2008)). However, either that category reduces to what agents are required to do or not. If the former, then it is thereby included in my discussion. If the latter, then it is beyond the scope of my discussion. Either way, I won’t spent time here on specific proposals of that kind.

This paper criticizes strong internalism about epistemic deontology by arguing that it is inconsistent with a promising response to a more general problem with epistemic deontology. Specifically, William Alston famously defended strong doxastic involuntarism—the view that very few of our doxastic attitudes are under our control. Given a voluntarist conception of responsibility, on which responsibility ultimately requires control, Alston’s view implies that epistemic deontology is false. A promising response to Alston’s view is to endorse compatibilism about doxastic attitude control. One specific version of compatibilism, reason-responsive compatibilism, is particularly promising for not only undermining Alston’s argument for strong doxastic involuntarism but also showing that many of our doxastic attitudes are under our control. However, I’ll argue that even if compatibilism is true—specifically, reason-responsive compatibilism is true—then strong internalism about epistemic deontology is false. But Alston was correct that strong doxastic involuntarism follows given his incompatibilist account of control. Thus, strong internalists cannot make use of either Alston’s incompatibilist account of control or a reason-responsive compatibilist account of control. Consequently, so long as we retain a voluntarist conception of responsibility on which responsibility requires control, we have reason for rejecting strong internalism about epistemic deontology.

In section I, I review Alston’s well-known defense of strong doxastic involuntarism and how it poses a problem for epistemic deontology. In section II, following Matthias Steup, I develop the compatibilist’s rescue package to Alston’s views, showing how reason-responsive compatibilism offers a promising response. In section III, I briefly defend that view from some recent criticisms of Meg Schmitt. In section IV, I exposit strong internalism about epistemic deontology, along with several particular internalist principles, before showing how they are inconsistent with reason-responsive compatibilism. Finally, in section V, I consider a possible “fall back position” strong internalists might be interested in. I concede that the position is immune to my criticisms, but maintain that it fails to retain a key motivation for strong internalism about epistemic deontology deriving from reflection on the first-person perspective.

I. Alston and Strong Doxastic Involuntarism

In a series of papers (1989), William Alston tried his hand at a well-known epistemological task: providing an account of epistemic justification. In so doing, Alston criticized what he called the “deontological” conception of justification. Proponents of the deontological conception of justification maintain, first and foremost, that our doxastic attitudes—beliefs, disbeliefs, withholdings—are subject to epistemic duties or obligations. They then propose accounts of justified beliefs in terms of such obligations or duties. For instance, a natural proposal is that a person is justified in adopting one of the three doxastic attitudes only if that person, in adopting that doxastic attitude, is not violating any of the epistemic obligations or duties her attitude is subject to (cf. Alston (1989: 86, 91)). But there are other ways. An alternative proposal might be that a person’s doxastic attitude is justified only if either the attitude is not in violation of any of the duties it is subject to or it is violation of some duties but the agent is blameless in violating them. To be clear, the deontological conception does not, by itself, specify what one’s epistemic obligations or duties *are*. It merely councils that our attitudes are subject to such duties, and that we understand epistemic justification in terms of them; such council was well received, as many epistemologists have thought of justification in this way.⁴

⁴ An alternative construal of the deontological conception of justification would be that a belief is justified if and only if it is not the case that it is in violation of a doxastic duty or obligation. However, it is conceptually possible that a belief is not in violation of a doxastic duty or obligation because none apply to it. (Compare: it’s not the case that I’m in violation of a company policy if I do not work for the company!) My understanding of the

However, Alston himself was no friend of the deontological conception of justification, and offered a formidable critique of it. At first pass, we can formulate his argument as:

- (1) Some proportion P of our beliefs, including perceptual, memorial, and testimonial beliefs, are justified.
- (2) If the deontological conception of justification is true, then any belief that is justified is subject to epistemic duties or obligations.
- (3) Thus, if the deontological conception of justification is true, then at least proportion P of our beliefs are subject to epistemic duties or obligations.
- (4) If at least proportion P of our beliefs are subject to epistemic duties or obligations, then at least proportion P of our beliefs are under our voluntary control.
- (5) The proportion of our beliefs under our voluntary control is less than P.
- (6) Therefore, the deontological conception of justification is false.

The first premise doesn't specify which proportion of our beliefs are justified, but presumably it is a high number. After all, a theory of justification with the result that few of our perceptual, memorial, or testimonial beliefs are justified is an inadequate theory of justification. The second premise spells out the deontological conception of justification. The fourth premise sees Alston endorsing a broadly voluntarist conception of responsibility, on which what an agent ought to do is something the agent can do. And while there are different ways of refining a "ought" implies "can" principle, like Alston I find myself convincing that some version of the principle is correct.⁵ Alston's defense of the fifth premise derives chiefly from a well-known passage in which he writes:

My argument ... simply consists in asking you to consider whether you have any such powers [to take up propositional attitudes at will]. Can you, at this moment, start to believe the United States is still a colony of Great Britain, just by deciding to do so? If you find it too incredible that you should be sufficiently motivated to try to believe this, suppose that someone offers you \$500,000,000 to believe it, and you are much more interested in the money than believing the truth. ... Can you switch propositional attitudes toward that proposition just by deciding to do so? It seems clear to me that I have no such power. (1989: 122) (cf. Plantinga, (1993: 24), Wolterstorff (1996: 112), Goldman (1999: 234))

Now I formulated Alston's argument in terms of "voluntary control" but the type of control that Alston is referring to in this passage is what he calls "basic voluntary control" (1989: 122-3)—the kind of control we have, for instance, over our limbs or thoughts. But there are other kinds of voluntary control that are not basic. (Alston calls these "indirect voluntary control" (1989: 127-52).) Further, as Richard Feldman (2000) and Brian Weatherson (2008) have pointed out, we *do* have indirect voluntary control over our beliefs. I have indirect control over beliefs about my immediate environment because I have control over my immediate environment. By opening the window, I will cease to believe that the window is closed. Perhaps some of our inferentially held beliefs are also under our long-term control in virtue of being able to reexamine the reasons for which we hold them. Perhaps, after some study, a philosopher might

deontological conception follows Alston in requiring that a belief is justified only if there are some set of doxastic duties or obligations that do apply to it.

⁵ Some philosophers simply deny this principle in response to Alston's argument (see, e.g., Feldman (2000), Schmitt (1992)). Others might deny it by offering up an alternative conception of responsibility that might be dubbed a "rationalist" conception instead of a "voluntarist." For that kind of approach to responsibility, see (e.g.) Adams (1985), Smith (2005), Hieronymi (2008). Such positions strike me as implausible. But I won't argue against them here as my concern is with a different response to Alston's argument.

come to revise his original view that Quine undermined the analytic/synthetic distinction. Such a belief might be said to be under control, insofar as there was a large set of actions that was undertaken that resulted in the philosopher going from belief to disbelief. However, these points do not undermine Alston's argument. For it is clear that the proportion of beliefs that are under my control in these other senses is lower than the proportion of beliefs that are justified. Thus, it seems, retreating to other kinds of voluntary control besides basic voluntary control provides little shelter in the face of Alston's argument.

This sub-argument of Alston's is intended to support what we might call "strong doxastic involuntarism," the position that the vast majority of our beliefs and degrees of belief are not under our voluntary control. We can separate Alston's strong doxastic involuntarism into further theses. First, almost none of our beliefs are under direct voluntary control. Second, that while some of our beliefs are under various kinds of indirect voluntary control, those beliefs are few and far between. Together, these theses support the position that the vast majority of our beliefs and degrees of belief are not under our voluntary control.

Alston used his strong doxastic involuntarism to criticize the deontological conception of justification. But clearly his argument for strong doxastic involuntarism is relevant, first and foremost, to whether epistemic deontology is true or not. Thus, his strong doxastic involuntarism is relevant to any position on which agents' attitudes are subject to epistemic obligations, regardless of the relation between those obligations and justification. And many philosophers are sympathetic to the idea that there are various epistemic obligations cognizers are subject to. To cherry pick two examples, many philosophers are sympathetic to the idea that in cases of (perceived) peer-disagreement subjects have an epistemic requirement to revise their beliefs or degrees of belief. The main dispute is simply exactly *how* agents ought to revise beliefs and degrees of belief (cf. Kelly (2005), Elga (2007), Feldman (2006)). Likewise, in exchanges of testimony, perhaps subjects have a requirement to not give a "deficit of credit" to a testifier on the basis of that person's inclusion in a certain social category (cf. Fricker (2007)). Perhaps we also have an obligation to not give an "excess of credit" to testifiers on the basis of their inclusion in certain social categories (cf. Davis (2016)). Clearly, these kinds of positions will also run afoul of Alston's strong doxastic involuntarism. For often times it is simply beyond our control to change our beliefs or degrees of belief in light of peer disagreement; likewise, how much credit a person gives a testifier is frequently outside their control. Thus, one could mount Alston-inspired arguments against these positions as well.

In short, the linchpin of Alston's argument is that strong doxastic involuntarism is true: very few of our beliefs or degrees of belief are under our control. Given a voluntarist conception of responsibility, he then used that claim to argue against the deontic conception of justification. But the same claim could be used to argue against a range of other versions of epistemic deontology that maintains that our beliefs or degrees of belief are subject to deontic appraisal.

II. The Compatibilist Rescue Package

The compatibilist about doxastic attitudes and degrees of belief (hereafter, just 'the compatibilist') offers a rescue package in the face of Alston and Alston-inspired arguments. This rescue package contains two parts. First, it alleges that Alston's argument for strong doxastic involuntarism fails. Second, it argues that given a suitably developed compatibilist view many, perhaps even most, of our beliefs are under our control.

A. Compatibilism Undercutting Alston's Doxastic Involuntarism

According to the compatibilist's response, Alston assumes an incompatibilist account of voluntary control in arguing that few of our beliefs are under voluntary control. Alston, recall,

asks us to *switch* doxastic attitudes—from disbelief to belief—regarding the proposition that the United States is a colony of Great Britain. But most of us are unable to do this. The most plausible explanation of our inability is not that we *fail* to execute a decision of ours—in the same way that a prisoner might *fail* to execute a decision to leave the prison because he finds his cell-door locked tight. Rather, the more plausible explanation is that we cannot even begin to decide to believe the United States is a colony of Great Britain, *given* the evidence we have against such a claim (cf. Steup (2011: 542ff.)).

On Alston’s view, in order for us to have control over an attitude, we must be able to decide to have that belief or not in the same circumstance (see, esp., (1989: 120-3)). But in the very same circumstance our evidence frequently keeps us from being able to decide to switch doxastic attitudes. Thus, on Alston’s view of control, those attitudes are not under our control. And, given the large number of attitudes we are unable to decide to change given the evidence, Alston’s strong doxastic involuntarism follows. However, compatibilists reject Alston’s view of control. Specifically, they reject that control requires being able to decide or not decide in the very same circumstances. Thus, if compatibilism is correct, then Alston’s argument for strong doxastic involuntarism fails.

B. Compatibilism, Reason-Responsive Compatibilism, and Control

But, the compatibilist continues that given a suitably developed compatibilist account many of our beliefs would also be under our control. Here the compatibilist does not challenge Alston’s claims about indirect voluntary control; she can, and should, agree that not many of our beliefs are under that kind of control. Rather, she maintains that many of our beliefs are under direct voluntary control, contrary to Alston’s claims.

In general, the compatibilist view is that whether or not our doxastic attitudes are under voluntary control turns on whether the causal processes in which they originated are good or bad. This naturally raises the question of which causal processes are “good” and which ones are “bad.” On this point, Matthias Steup writes,

No matter *how* the distinction between good and bad causal processes is cashed out, surely we should agree, to begin with, that the standard origins of belief—perception, introspection, memory, reasoning—are paradigmatically good causes. A version of compatibilism that classifies our standard belief-forming processes as bad, thus equating them with mental illness and the like, would be unacceptable. (2011: 548)

Steup has gone on to illustrate different compatibilists accounts and how, on each, beliefs can be considered voluntary (2000, 2008). I won’t take up each of them. Instead, I’ll follow Steup in providing what he and I regard as a particular promising way of spelling out a compatibilist notion of “good causes” in terms of causes that are “reason responsive” (2008: 379-381; 2011: 547-8). This notion gets its most well-known articulation by Fischer and Ravizza (1998). At first pass, their idea can be put like this:

An agent’s decision or choice to X is caused in a reason responsive way iff in nearby worlds the agent tends to decide or choose to X if and only if the totality of the agent’s reasons support Xing.⁶

⁶ Compare Steup (2011: 547). This is an approximate gloss on what Fischer and Ravizza refer to as “moderate reason-responsiveness.” Roughly, this is a *moderate* requirement because it is stronger than the *weak* requirement that there only be one such world where the agent is reason-responsive but it is weaker than the *strong* requirement that in every such world the agent is reason-responsive. See Fischer and Ravizza (1998: 41ff.) and Fischer (2006)

Now this account of reason responsiveness only applies to decisions or choices, and what we need is an account that applies to doxastic attitudes. An analogous principle would be:

An agent's doxastic attitude X is caused/maintained in a reason responsive way iff in nearby worlds the agent tends to adopt attitude X if and only if the totality of the agent's reasons support adopting X.

On this proposal, then, an agent's doxastic attitudes are under her control when they are caused in a reason responsive way; further, *contra* Alston, being caused in a reason-responsive way does not require an ability to be able to decide to adopt or not adopt the attitude in the same circumstance.

By itself, this principle does not imply that any of our doxastic attitudes are actually caused in a reason responsive way and thus under our control. But it is natural to fill in an account of reasons on which many of our beliefs are caused in a reason responsive way. For instance, suppose we plausibly assume that evidence, and some experiences as well, furnish reasons. Many of our beliefs are formed in a way that is responsive to evidence and experiences. For instance, I believe I have hands, disbelieve the earth has no moons, and withhold on the number of black footed ferrets is even. These attitudes are reason responsive. If my reasons were to change, so would my attitudes. For instance, if I were to wake up one day with bandages at my elbows and nothing beyond, my totality of reasons would not support that I have hands and I would not believe that I have hands. More generally, many of our doxastic attitudes tend to vary, modally, with our reasons and that is sufficient for them to be caused in a reason-responsive way.

Of course, there are other compatibilist accounts of control beyond reason-responsive compatibilism. These other accounts are rivals to reason-responsive compatibilism. However, I focus on reason-responsive compatibilism specifically for several reasons.

First, other versions of compatibilism may be unhelpful to proponents of epistemic deontology, such as proponents of the deontological conception of justification. The main problem other versions of compatibilism might face is that they require too much cognitive sophistication to be plausible when applied to doxastic attitudes. For instance, other compatibilist accounts may require for control that agents' desires play an important role.⁷ A simple account may be that an agent's doxastic attitude is under her control only if she wants to have that attitude. More complicated accounts might be that a doxastic attitude is under her control only if she has it *because* she wants it. And, of course, even more complicated versions are possible. An account inspired by Frankfurt's work (1971), might maintain that a doxastic attitude is under an agent's control when it is caused by a set of her desires that not only move her to form the attitude but are also desires that she has a second-order desire to be the ones that cause her attitudes.

These kinds of accounts will not imply that many of our doxastic attitudes are under our control. For it is doubtful that many of our beliefs are caused by our desires; it is even less likely that most agents have second-order desires about their first-order desires about attitude formation that would allow this account to apply to enough of an agent's beliefs as needed by various proponents of epistemic deontology. Since reason-responsive compatibilism need not require such cognitive sophistication for control, it is more plausible as an account of belief control that proponents of epistemic deontology need for their position.

for details and elaborations on "reason-responsiveness" and the differences between strong, moderate, and weak reason-responsiveness.

⁷ Steup (2008) canvasses a range of non-reason responsive compatibilist positions and many of them give desires and wants an important role.

Second, reason-responsive compatibilism naturally fits many ideas that epistemologists are already sympathetic to. Compatibilists about doxastic attitudes will need to mark a distinction between “good” sources of belief that give rise to control and “bad” sources of belief that do not. They will further maintain that certain sources of belief—perception, memory, introspection, etc.—are good sources of belief whereas other—biases, neuroses, etc.—are bad sources of belief. But many epistemologists already believe that there are reasons for belief and that our cognitive attitudes can be responsive to them. They also normally believe that perception, memory, introspection, etc. are normally responsive to reasons and biases, neuroses, etc. are not responsive to reasons. Thus, epistemologists are already separating sources of belief in close to the same way as the compatibilist will want. Additionally, epistemologists are separating these sources of belief in these ways partly because of the way these sources are reason-responsive (or not). It is thus natural for compatibilists to maintain that what makes those sources of belief—perception, memory, introspection, etc.—good ones that give rise to control is their reason-responsiveness and what makes other sources of belief—biases, neuroses, etc.—bad ones that do not give rise to control is their lack of reason-responsiveness. But this is just what reason-responsive compatibilism maintains. In this way, reason-responsive compatibilism naturally fits ideas that epistemologists are already sympathetic to.

Third, because it is plausible that many of our beliefs are formed in reason responsive ways, reason-responsive compatibilism is well suited to respond to Alston’s argument. For if it is right, then many of our beliefs that originate in reason-responsive sources, like perception, memory, or testimony, will be under our voluntary control. Thus, premise (5) of Alston’s argument from above will be false.

Additionally, it is also possible that the reason-responsive compatibilist position could be further developed to help defend some of the other positions that are vulnerable to an Alston-style critique. To illustrate with the case of peer disagreement, suppose the fact that a perceived peer disagrees with me on p is a reason for me to revise my belief that p . If in most nearby worlds I tend to revise my belief just when confronted with reasons to revise my belief like peer disagreement, then that would be sufficient for showing that my belief is formed in a reason responsive way and thus under my control.

To be sure, none of this proves that if one is a compatibilist about doxastic attitudes, then one must be a reason-responsive compatibilist. For instance, one could just think it a coincidence, happy or otherwise, that good sources of belief are reason-responsive while maintaining that it is some other feature of those beliefs that explains why beliefs so formed are under control. And, I admit, there are other forms of compatibilism about actions that one might develop for compatibilism about doxastic attitudes that I have not considered. But it does indicate that if a proponent of epistemic deontology is to respond to Alston and Alston-inspired arguments by embracing a kind of compatibilism, reason-responsive compatibilism is particularly promising, if not the most promising approach.

Summing up, a compatibilist account of voluntary control—specifically, a reason-responsive version of compatibilism—offers a promising rescue package to Alston and Alston’s inspired arguments. It undermines Alston’s argument for strong doxastic involuntarism. Further, given some development, it provides a promising way to respond to Alston’s argument and Alston inspired arguments against various versions of epistemic deontology.

III. Schmitt’s Criticism of the Compatibilist Rescue Package

The compatibilist strategy—and specifically Steup’s particular brand—has come under criticism. Steup has responded to many of the criticisms (see, inter alia, his (2000, 2008, 2011,

2012, 2017); see also Ryan (2003)). But here I will respond to a pair of unanswered criticism due to Meg Schmitt (2015). Schmitt claims that (i) Steup fails to identify the compatibilist account of voluntary control more generally and (ii) the specific reason-responsive proposal fails to secure voluntary control.

As we just saw, Steup claims that we can meaningfully draw a distinction between good sources of belief and bad sources without necessarily endorsing a particular account of the difference, even if he does show affinity for the reason-responsiveness approach. Schmitt objects to that claim. Commenting on the above passage from Steup, she writes,

Even if this is correct and our beliefs typically result from the aforementioned causes and these causes should be considered ‘good’ on whatever account we advance, surely the same line of reasoning could apply to any number of states associated with the proper function of a human being. But this entails that a wide range of states that are not *in fact* exercises of volitional control are exercises of volitional control. (2015: 29)

Schmitt illustrates this point with an example involving eye color. Most human beings have the eye color they do because of the result of their parents’ genetic makeup and not (e.g.) physical manipulation or radiation poisoning. Thus, insofar as one’s eye colors are a result of processes that are properly functioning, it implies that one’s eye color could also be under voluntary control. But this is an absurd result.

Schmitt represents this criticism as showing that Steup has failed to identify the compatibilist’s criteria for voluntary control in terms of “good” sources (2015: 26, 28). I think that’s mistaken. A more precise way of understanding this criticism is that it shows that Steup is wrong in claiming that *any* way of distinguishing between “good” and “bad” causes will secure the result that the compatibilist wants. For one way of cashing out the idea of a good cause is in terms of the *proper functioning* of human beings, which is what Schmitt does in the above passage. But cashing out the idea of good cause in *that* way will imply that too much is under our voluntary control. So the compatibilist may retain the idea that voluntary control tracks “good” causes, but simply hold that not just any account of a “good” cause will do.

Turning to the specific reason-responsive account of voluntary control, Schmitt objects that the account overextends (2015: 30). Her objection rests on the fact that ‘reason’ can be used in (what she calls) a “merely causal sense.” For instance, one might say “the reason why the forest is on fire is because the volcano erupted.” Her point is that many things which are not under one’s voluntary control—such as one’s eye color—are the result of “reason responsive” mechanisms, if we understand ‘reason’ here in a merely causal sense. To avoid this overextending, the compatibilist needs to identify a sense of ‘reason’ that is both different from ‘reason’ in this merely causal sense *and* can underlie voluntary control for both action and belief. She argues that it is hard to see how to do this (2015: 32ff.).

The compatibilist can overcome this challenge. As some philosophers have pointed out (e.g. Scanlon (1998), Parfit (2011)), we can identify a “normative” usage of the term ‘reason’ that is distinct from the merely causal sense of ‘reason.’ Of course, Schmitt is right that merely *claiming* that there is a difference does not *explain* the difference between the two (2015: 33). But we can explain the difference by pointing out the connections between normative reasons and other normative concepts. Specifically, when something is a normative reason, it is a consideration in favor of something. For instance, if a reason to perform an action is that I can prevent needless harm, then that is a consideration in favor of performing that action. By contrast, it need not be the case that merely causal reasons are considerations in favor of things.

Perhaps there are considerations in favor of a particular forest fire occurring. (For instance, forest fires can reset ecosystems, and perhaps that is valuable.) But it does not follow from the fact that a forest fire was caused by an erupting volcano, that the volcano erupting was a consideration in favor of the forest fire. Thus, we can explain the difference between normative reasons and merely causal reasons by pointing out that normative reasons connect to other normative concepts, like being a consideration in favor of, and that is not necessarily the case with merely causal reasons.⁸

This normative sense of ‘reason’ can be used by the compatibilist. She can maintain that a doxastic attitude is under voluntary control when it is formed or maintained by a belief forming mechanism that is responsive to normative reasons. Since normative reasons are not identical to merely causal reasons, Schmitt’s argument fails to show that compatibilist accounts overextend. Of course, more could be said to more fully developing the distinction between normative reasons and merely causal reasons. But they do offer some promising resources for responding to Schmitt’s criticisms. However, the appeal to compatibilism—specifically, reason-responsive compatibilism—comes at a cost, to which I now turn.

IV. The Cost of Compatibilism

Though compatibilism might provide a promising response to Alston-inspired arguments, I argue that this response comes at a cost. Specifically, even if reason-responsive compatibilism for doxastic attitudes is true, then there are theories of epistemic obligations or epistemic “oughts” that are false. Specifically, a cluster of theories I label “strong internalist” will conflict with reason-responsive compatibilism, given a voluntarist account of responsibility. Thus, responding to Alston-inspired arguments by appealing to compatibilism in this particular way comes at a cost.

A. *Internalism about Epistemic Obligations*

Internalism in epistemology is a cluster of views. In general, what unifies these views is that they propose accounts of properties that place an important role on what an agent can access immediately, introspectively, or via reflection.⁹ More specifically, what I’ll call “weak” internalism maintains that if an agent’s doxastic attitude has P, then some of the facts that determine or “fix” that her attitude has P are facts that she can access. By contrast, “strong” internalism maintains that if an agent’s doxastic attitude has P, then there are no facts that determine or “fix” that her attitude has P that she cannot access. To put these positions in terms of supervenience, weak internalism about P maintains that part of the supervenience base contains facts that an agent can access; strong internalism maintains that none of the supervenience base contains facts that she cannot access. The guiding idea behind strong

⁸ Of course, more could be said about the relation of ‘consideration in favor of.’ Some, like Scanlon, understand it in terms of the concept ‘being a reason,’ generating a tight circle of primitive concepts; but there are other understandings as well. But the foregoing does show that Schmitt is incorrect when she claims that “describing a reason as ‘normative’ seems nothing much more than describing it as ‘non-explanatory’” (2015:33).

⁹ For more on access and internalism, see Schmitt (1992), BonJour (2002), Bergmann (2006). For an alternative proposal of internalism in terms of supervenience and mental states, see Conee and Feldman (2001). For criticism of that proposal, see Bergmann (2006). Some authors (e.g., Bergmann (2006)) formulate “internalism” about justification as a version of what I’ve called “weak internalism”, while others (e.g., BonJour (2002)) formulate it as a version of what I’ve called “strong internalism.” What an agent can access might include more than just her mental states. It may also include general epistemic principles (cf. Chisholm (1966)) or logical/probabilistic principles (cf. BonJour (2002)). The inclusion of “accessible” general principles does not matter for my argument. Finally, keep in mind that internalists will want to avoid as “accessible” information that is gained through further research, e.g., consulting an expert or online encyclopedia. This is to stay true to the idea that things outside the perspective of an agent are irrelevant. I assume we have a workable notion of “access” on which that is true.

internalism about some property P is that facts *outside* the purview of the agent are irrelevant to whether her attitudes have P. In this section, I will be concerned with a version of strong internalism; I'll briefly discuss a version of weak internalism in the next section.

The strong internalist position I'm concerned with is:

Strong Epistemic Internalism about Epistemic Oughts: What doxastic attitudes or degrees of belief a subject ought to have, ought not have, or are permitted to have are determined just by facts the agent can access; alternatively, what doxastic attitudes or degrees of belief a subject ought to have, ought not have, or are permitted to have supervenes only on facts that she can access.

(Hereafter, *Strong Epistemic Internalism*.) This position is a form of “strong” internalism, since it holds that no facts outside the access of an agent determine the deontic properties of that agent's doxastic attitudes.

Some authors propose principles that specify sufficient conditions for what doxastic attitudes agents ought to have. Sometimes those conditions are such that agents can always access when those conditions are met. I will refer to these as “strong internalist principles.” They do not imply *Strong Epistemic Internalism*, since they only specify sufficient conditions. Nor does *Strong Epistemic Internalism* imply them, since it does not specify sufficient conditions for when agents ought to have or not have a doxastic attitude. Nonetheless, they are consistent with *Strong Epistemic Internalism* and some of them might be understood as substantive proposals in filling out the idea of *Strong Epistemic Internalism* that at least some of the time what goes on outside an agent's purview is irrelevant to what she ought to believe.

Here are some examples of strong internalist principles. Consider:

The Seemings Principle. If it seems to you that *p*, and you lack any defeaters for believing that *p*, then you ought to believe that *p*. (cf. Huemer (2007, 2013), Madison (2017); perhaps Tucker (2010))

(Defeaters are other beliefs or experience that one can access which draw into question the propriety of one believing that *p*.) Though there are various disputes about how best to understand what seemings are (see Moretti (2015) for an overview), there appears to be agreement that when something seems to you to be a certain way, you are occurrently aware of that fact. Thus, *The Seemings Principle* is a strong internalist principle.

Consider also the following position:

Epistemic Conservativism: If a subject believes that *p* and lacks any defeaters for that belief, then the subject is permitted in believing *p*.

This principle, or something close to it, can be found in the work of various philosophers including perhaps Quine (1960), Sklar (1975), Chisholm (1980), Harman (1986), Elgin (1996), McGrath (2007), McCain (2008), Poston (2014).¹⁰ For reasons immaterial to discussion here, it is not always clear to me that whether an agent believes something is accessible to that agent.

But we can consider a logically weaker position than *Epistemic Conservativism*:

Weaker Epistemic Conservativism: If a subject believes that *p*, can access that she believes that *p*, and lacks any defeaters for that belief, then the subject is permitted in believing *p*.

¹⁰ Not all of these authors use the term ‘permitted’ and some endorse a weaker position than the one formulated here. For reasons I cannot get into here, I think authors who prefer a weaker formulation should opt for this stronger formulation. In any case, I'm concerned with the principle as formulated, not necessarily the detailed exegesis of all of these authors. For additional discussions of epistemic conservatism, see Kvanvig (1989), Christensen (1994), and Vahid (2004).

Weaker Epistemic Conservatism is a strong internalist principle.

A final example. The evidentialist Richard Feldman has proposed the following principle:
The Evidence Principle: You ought to adopt the doxastic attitude—belief, disbelief, suspended belief—towards *p* that fits the evidence you have. (Feldman (2000, 2002))

The *Evidence Principle* needs to be filled in by giving an account of three things: evidence, “had” or possessed evidence, and “fit.” Some evidentialists give an account of evidence on which it consists of non-factive, representational mental states (Feldman and Conee (2004), McCain (2014)). Evidentialists give accounts of evidence possession in terms of the occurrent mental states one is aware of (Feldman (1988a)) or those states as well as the ones accessible upon reflection (McCain (2014)).¹¹ Those accounts imply that the evidence one has are mental states that one can access. Evidentialists are usually less forthcoming with an account of “evidential fit.” Presumably, that concept can be reduced to the concept of evidential support. A belief that *p* “fits” evidence when *p* is supported by it; disbelief that *p* “fits” evidence when the negation of *p* is supported by it; withholding that *p* “fits” the evidence when it neither supports *p* nor the negation of *p* (cf. Conee and Feldman (2008: 83)). In any case, I’ll assume evidential support in such a way that agents can access whether a doxastic attitude is supported by their evidence. This is an assumption some evidentialist should be congenial to.¹² Specifying the *Evidence Principle* these ways results in a principle that is a version of *Strong Epistemic Internalism*.¹³

A final comment on these principles. Several of these principles specify sufficient conditions for what an agent ought to believe. Some might object that these principles are too strong (cf. Nelson (2010)). For instance, these principles will imply that even if an agent is uninterested in a proposition, in certain circumstances, she is nonetheless obligated to form an attitude about that proposition. Some might find that consequence too strong. We can sidestep this issue by simply reformulating those principles not in terms of what agents’ ought to believe but what they are permitted to believe. However, the difference between formulations in terms of what agents ought to believe or are permitted to believe won’t matter for my discussion, so I’ll ignore this complication.

B. A Problem for Strong Epistemic Internalism

Crudely put, the problem for *Strong Epistemic Internalism* is this. *Strong Epistemic Internalism* is a version of epistemic deontology and a general problem with epistemic deontology is to address Alston-inspired arguments. So long as we are incompatibilists about control, then Alston-inspired arguments win the day: our doxastic attitudes are not under our control and *Strong Epistemic Internalism* is false. However, suppose we are compatibilists about control. Then the most promising version of compatibilism for proponents of epistemic deontology is a version of reason-responsive compatibilism. However, even if reason-responsive compatibilism is true for doxastic attitudes, *Strong Epistemic Internalism* is false. Thus, while embracing reason-responsive compatibilism provides a promising response to Alston-inspired arguments, it comes at a cost: *Strong Epistemic Internalism* is false.

¹¹ For more on evidence possession and evidentialism, see my (under review).

¹² For instance, both Conee and Feldman (2008) and McCain (2014) suggest accounts of evidential fit on which, crudely, evidential fit of a proposition to body of evidence is a matter of explanatory relations between them available to subjects. But an explanation being available to a subject is defined in a way that implies it is accessible to a subject.

¹³ Feldman is also a mentalist about justification: he thinks that any two subjects that are alike “mentally” are alike in terms of justification. But I’m not here focused on his theory of justification but this principle he articulates.

Of course, this presentation *is* a crude presentation. It needs to be qualified in certain ways. For instance, some of our beliefs are under our control in virtue of being under long term indirect voluntary control. However, not enough are under our control as required by various proponents of *Strong Epistemic Internalism*. Additionally, this presentation does contain several lacunas. It assumes that responsibility requires control. It also assumes that the most promising version of compatibilism about doxastic attitudes is a reason-responsive version of compatibilism. Some of these assumptions I have already defended here. Instead of offering further defenses of them, I'll turn my attention to what is the most surprising claim here: that even if reason-responsive compatibilism is true for doxastic attitudes, *Strong Epistemic Internalism* is false.

If reason-responsive compatibilism is true, then one's doxastic attitude is under one's control only if it originates in a reason-responsive mechanism. For it to originate in a reason-responsive mechanism that attitude must tend to vary, modally, with one's reasons. Thus, whether one's doxastic attitude is under one's control is fixed or determined not just by how *as a matter of fact* one came to acquire that attitude but also by how one *would* acquire such an attitude as well.

However, these various strong internalist principles imply that what doxastic attitudes an agent ought to have or is permitted to have are fixed or determined only by *actual* facts, specifically, the actual facts that agents can access. But, the actual facts that an agent can access need not, and usually do not, settle how an agent would acquire their doxastic attitudes in certain counterfactual scenarios. Consequently, it will be possible to come up with cases where these various strong internalist principles imply that there is some doxastic attitude an agent ought have or is permitted to have where, given reason-responsive compatibilism, that attitude is not under the agent's control. Given a voluntarist conception of responsibility, it follows that such attitudes are not open to deontic evaluation *contra* these various strong internalist principles. Thus, if reason-responsive compatibilism is true, then these various strong internalist principles are false.

An example will make this point more intuitive. Suppose Julius is staring at a screen that is about to flash simple geometric shapes. The screen flashes a square and it seems to Julius that it flashes a square. He has no defeaters indicating otherwise. He forms the belief that the figure is in fact a square. Given the *Seemings Principle*, it follows that Julius ought to believe that the figure is a square. However, suppose that unbeknownst to Julius, his cognitive mechanisms are not reason-responsive: even if things were to have seemed to him differently—if the shape were to have seemed to be (e.g.) a circle or trapezoid—he still would have formed the belief that it was a square. Given reason-responsive compatibilism, his doxastic attitude is not under his control and so ineligible to be evaluated as one he ought to hold *contra* the *Seemings Principle*.

Though I've focused on the *Seemings Principle*, it is easy to see the recipe for generating conflicts between reason-responsive compatibilism and the other principles. Describe a case in which a subject's doxastic attitudes meet the sufficient conditions specified by the principle. Given that principle, it will follow the subject's doxastic attitude has some deontic property. Now add further details to the scenario indicating that in nearby possible worlds that agent's doxastic attitude say the same while her reasons change. This will be possible, since these principles place no constraints as to what goes on in nearby possible worlds. The resulting scenario will be one in which an internalist principle implies the attitude has some deontic property, but reason-responsive compatibilism implies it does not. Thus, if reason-responsive compatibilism is true, the various strong internalist principles are false.

In response, proponents of these internalist principles might point out that agents can access their memory upon reflection. Further, they might claim that an agent's memory might provide evidence that an agent's cognitive faculties are reason-responsive. Indeed, insofar as many of us have reasons for thinking our cognitive faculties are reason-responsive, it is because of their track record.

Even if these points are correct, they fail to undermine my argument for several reasons. First, and foremost, all my argument requires is that an agent's faculties are *not* reason-responsive. My argument does not and need not make any claims about whether an agent has any reason to believe that her faculties are reason-responsive. Further, it is clearly possible that an agent's faculties are not reason-responsive and yet she has memories indicating that they are. (For instance, perhaps the relevant faculties *were* reason-responsive, but have ceased to be.) Second, having a memory that X which one can access is not sufficient for having access to X. I might have a collection of memories that my car is reliable; but that does not mean I have introspective access to the reliability of my car! So pointing out that people often times have memories which would support that their cognitive faculties are reason-responsive does not undermine my argument.

It might be thought that these various principles can be salvaged by incorporating claims about how agents "base" their doxastic attitudes. To be sure, this kind of response is really to reject the early principles but accept an amended version of them that somehow incorporates claims about basing, e.g.,:

Base-Amended Seemings Principle. If it seems to you that p , you lack any defeaters for believing that p , and you base your belief that p on your seemings that p , then you ought to believe that p .

However, even such amended versions will not help avoid the problem. For instance, to generate a conflict between *Base-Amended Seemings Principle* and reason-responsive compatibilism, we only need to add a further claim to our earlier case involving Julius. Specifically, all we need add is that Julius' belief that the figure is a square is based on her seemings that it is a square. Then *Base-Amended Seemings Principle* will imply that Julius ought to believe that the figure is a square, while reason-responsive compatibilism implies the opposite.¹⁴

Central to my argument is the following idea. We can keep the same what facts an agent can introspectively access, while providing cases that vary the reason-responsiveness of the mechanisms that give rise to agent's doxastic attitudes.¹⁵ But belief forming mechanisms are plausibly thought of as clusters of dispositions. Further, it is these clusters of dispositions that can (partly) explain how an agent forms doxastic attitudes in nearby possible worlds. Thus, one possible response proponents of *Strong Internalism* might propose is that, as a matter of fact,

¹⁴ One might object that, depending on the account of basing, adding the further claim that Julius' belief is based on her seemings and is not reason-responsive actually produces an incoherent scenario. This could only happen if there were an account of the basing relation on which, when a doxastic attitude is based on an agent's reasons, it was impossible for the attitude to not be reason-responsive. I'm not sure there is a plausible account of the basing relation that would entail that. But even if there were, it could not be used by internalists to respond to this problem. Since being reason-responsive is not determined by the actual facts that agents can always access, and this account of basing relation implies being reason-responsive, it follows that agents could not always access how their doxastic attitudes are based. Thus, amended principles that used this account of the basing relation will not be internalist ones.

¹⁵ It does not matter if the further details amount to either changing the reason-responsiveness of the mechanism that give rise to the agent's attitudes in the actual world or actually changing the mechanism in the actual world from a reason-responsive one to a non-reason responsive one. Either way will result in a scenario that is consistent with these strong internalist principles and *Strong Epistemic Internalism*.

agents can introspectively access their belief forming dispositions. Further, if they could, agents could introspectively access whether their doxastic attitudes are open to deontic appraisals. This conflicts with my statement that we can keep the same introspectively accessible facts, while varying facts about reason-responsiveness.

At a conceptual level, there is nothing problematic with this response. That is, there is no conceptual barrier to saying that a belief is caused by a reason-responsive mechanism and it is introspectively accessible that the mechanism is reason-responsive. The problem with this response is empirical. For it is simply false that agents have introspective access to many, even most, of the dispositions and cognitive mechanisms that are responsible for the forming and retaining of our doxastic attitudes. (See Kornblith (2002, 2012) for a helpful overview of some of this literature.) Thus, it will be possible to keep the same facts about what an agent can introspectively access, while varying the reason responsiveness of the mechanisms that give rise to the agent's attitudes.¹⁶

The immediate upshot of the argument of the last few pages is that there is a conflict between *Strong Epistemic Internalism*—and various strong internalist principles—and reason-responsive compatibilism. Specifically, given a voluntarist conception of responsibility, then if reason-responsive compatibilism is true, *Strong Epistemic Internalism* is not. Reason-responsive compatibilism is only one articulation of compatibilism. But, I argued earlier, reason-responsive compatibilism is a particularly promising form of compatibilism both in general and for responding to Alston-like arguments. Thus, given a voluntarist conception of responsibility, we have reason for thinking that if compatibilism *simpliciter* is true, then *Strong Epistemic Internalism* is false. To be sure, that reason can be outweighed by articulating a compatibilist account of control that was (i) distinct from reason-responsible compatibilism, (ii) implied that many of our doxastic attitudes are under our control, and (iii) was consistent with *Strong Epistemic Internalism*. But until such an account is articulated, we have reason for thinking if compatibilism is true, then *Strong Epistemic Internalism* is false. And if incompatibilism is true, then *Strong Epistemic Internalism* is false. Thus, given a voluntarist conception of responsibility, we have reason for rejecting *Strong Epistemic Internalism*.

V. Introspection and Internalism

In response to this criticism, proponents of internalism may opt for what I'll call the *Hybrid Position*. This position marks a distinction between the conditions under which a doxastic attitude is open to deontic appraisal and the particular deontic appraisals of those attitudes. On the *Hybrid Position*, considerations outside the access of an agent can be relevant to whether a doxastic attitude is open to deontic appraisal. But given that the attitude is open to doxastic appraisal, the factors that determine the particular deontic appraisal are always accessible to the agent. Proponents of the *Hybrid Position* may even update the previous principles accordingly, perhaps like:

Hybrid-Modified Epistemic Conservativism: So long as the belief that *p* originates in or is maintained in a reason-responsive way, if a subject believes that *p* and lacks any defeaters for that belief, then the subject is permitted in believing *p*.

¹⁶ To be sure, as I discuss in (2014), we frequently know the crude origins of our beliefs (e.g., that I have this belief because of what I saw or because so-and-so told me). However, we do not retain the origins of all of our justified beliefs. Further, knowledge of crude origins can be insufficient for knowledge of reason-responsiveness because the particular mechanisms underwriting belief are not necessarily transparently accessible to us. For a helpful overview of some of the ways that biases and unreliable heuristics can influence simple reasoning in a way that is not transparent, see Kahneman (2011).

To be clear, the *Hybrid Position* requires rejecting *Strong Epistemic Internalism*. In fact, it amounts to a version of what we might call *Weak Epistemic Internalism*. However, it is a distinctive version of *Weak Epistemic Internalism*. For it refuses to give considerations outside an agent's access any role in determining the particular or specific deontic properties of doxastic attitudes. But the *Hybrid Position* as well as principles modified in line with it are immune to my criticisms of *Strong Epistemic Internalism*. Modified principles like *Hybrid-Modified Epistemic Conservatism* explicitly incorporate claims about reason-responsiveness. Further, the *Hybrid View* avoids the claim that *Strong Epistemic Internalism* makes that I objected to, namely, that no factor outside what an agent can access is relevant to deontic appraisal. What would be the drawback on such a response?¹⁷

The main drawback of this response is that a substantive motivation for *Strong Epistemic Internalism* fails to motivate the positions required by this response. The substantive motivation I'll call the *Regulative, First-Person Motivation*. That motivation supports *Strong Epistemic Internalism*. It could also be used by proponents of principles that are version of *Strong Epistemic Internalism*. But, as I'll now argue, it could not be used by proponents of the *Hybrid Position*.

The *Regulative, First-Person Motivation* derives from how a certain constraint on certain questions motivates *Strong Epistemic Internalism*. The questions at stake are questions like 'What ought I believe now?', 'Given my situation, which of my doxastic attitudes are the ones I ought to have?' or 'What beliefs am I required to retain? Which am I required to abandon?' Though each of these questions is slightly different, they form a close-knit of questions. These questions are ones that both internalists and externalists might regard as central. Further, these questions may not be without historical precedent. For instance, Descartes' work in epistemology, whether it be the published *Mediations* or the unpublished 'Rules for the Direction of the Mind,' is naturally interpreted as being concerned with such questions.

Those inclined toward *Strong Epistemic Internalism* are likely to propose a distinctive constraint on answering these questions. That constraint is that an adequate answer to these questions must be one that agents can use as a guide or instructions from a first-person, deliberative point of view (cf. Goldman (1980)). This constraint rules out certain kinds of answers. For instance, being told that one ought to conduct further research or seek out more evidence fails this constraint. For such answers do not tell me what I ought to believe *now* or what I ought to believe given my current epistemic situation (cf. Chisholm (1966: 24-5), Feldman (1988b: 250)). But even being told (e.g.) to believe what originated in a reliable source or retain the beliefs that agree with the experts will fail this constraint, since from an agent's first-person perspective she might be unsure which beliefs originated in a reliable source or be unable to tell what the experts think (cf. Foley (1993: 117ff.)). Rather, what agents asking these questions from a deliberative first-person perspective are requesting is a criteria or guide which they could employ without further empirical work and would be manifest to them from that perspective. Such a guide or instruction need not necessarily influence their beliefs, but could

¹⁷ Given the deontological conception of justification, we can distinguish between *Strong Epistemic Internalism* and the *Hybrid Position* using the terminology of "justifiers." Call a "global justifier" anything that is relevant to whether an agents' belief is justified. Call a "conferring justifier" anything that confers or robs justification to a belief. Clearly conferring justifiers are a subset of global justifiers, but not conversely. *Strong Epistemic Internalism* maintains that all global justifiers are accessible to an agent. *Hybrid Position* maintains that all conferring justifiers are accessible to an agent, but not all global justifiers may be.

simply play the role of identifying or marking the doxastic attitudes they ought to hold (cf. BonJour (2002)).

Accepting this constraint naturally pushes us to *Strong Epistemic Internalism*. For that position tells us that what facts determine what we ought to believe are accessible to agents via introspection and reflection. Thus, when an agent deliberates from her first-person perspective, she should be able to help regulate her beliefs. By contrast, if some fact played some role in determining the deontic status of an agent's doxastic attitudes, but was not accessible to her, the agent would be unable to regulate her beliefs from her first-person perspective. Since this motivation for *Strong Epistemic Internalism* derives from a constraint that emphasizes how an agent might regulate her beliefs from a first-person perspective, I call it the *Regulative, First-Person Motivation*.

However, this motivation does not support the *Hybrid Position*. For this motivation supports *Strong Epistemic Internalism* and not a weaker position. For a weaker position, like the *Hybrid Position*, would permit considerations outside the access of the agent to play a role in the deontic appraisal of an agent's doxastic attitudes which is exactly what this motivation is trying to avoid. To be sure, the *Hybrid Position* only permits inaccessible considerations to play a role in determining whether an agent's attitudes are open to deontic evaluation, and not the particular deontic properties of the agent's attitudes. But this distinction is unimportant from the first-person perspective, since to deliberately arrive at the conclusion that (e.g.) this belief is one that ought not be held, one must also arrive at the conclusion that it is appropriate to evaluate this belief as the sort of thing that is open to deontic evaluations to begin with. After all, the relevant questions at stake are not things like, "if my belief is under my control, should I maintain it?" but "should I maintain this belief?"¹⁸

To be sure, there might be other motivations for the *Hybrid Position*. Though I cannot prove it, I think it will be difficult to find some. For instance, some standard motivations for internalism about justification do not support the *Hybrid Position*. Perhaps the major motivation for internalism is some version of BonJour's clairvoyance argument (1985). But that argument, if successful, would only show that a necessary condition for justification is that one have access to some of what it is that justifies one's beliefs. It leaves open that factors outside the perspective of the agent could play a role in what an agent is justified in believing. Thus, such an argument would not establish that considerations outside the perspective of the agent could not play a role in the deontic status of an agent's beliefs, which is what the *Hybrid Position* maintains.

Some might reach for evil-demon thought experiments to support the *Hybrid Position*. An evil-demon thought experiment involves two agents who are qualitatively alike mentally, but one of which is being massively deceived by an evil demon. Intuitions are then "pumped" about

¹⁸ It might be suggested that subject treat these various internalist principles as general rules of thumb, with the agent treating it as a background assumption that his attitudes are open to deontic appraisal. I'm not sure this suggestion helps with the problem for two reasons. First, even if an agent did do this, it is still possible that such "general rules of thumb" would lead them astray, because the background assumption is false and the agent cannot introspectively tell that it is false. Second, this response would make it even harder to motivate the *Hybrid Position* because it would undermine its ability to criticize rival views. For instance, consider a view that endorsed a reliabilist constraint, on which an agent ought to believe something only if it originated in a reliable source. A defender of the *Hybrid Position* might try to object to such a constraint on the basis of the role inaccessible facts play in determining deontic appraisals. But proponents of that constraint could offer a response that mirrors this response to my criticism: the principle is a general rule of thumb, with agents treating it as a background assumption that their attitudes originate in reliable sources. If the above response succeeds in defending the *Hybrid Position*, I don't see why it wouldn't also succeed in defending this reliabilist constraint.

whether or not both subjects have some salient epistemic property P. (See, e.g., Cohen (1984), Conee and Feldman (2001), Silins (2005), and Lyons (2013) for more on such thought experiments.) However, such thoughts experience are usually thought to support, at best, “mentalism” about P: the view that two subjects can differ on whether their beliefs have P only if they differ “mentally.” Since we are concerned with deontic properties of beliefs, the relevant version of mentalism here is *Deontic Mentalism*: the doxastic attitudes of two agents can differ in terms of their deontic properties only if they differ “mentally.”

However, *Deontic Mentalism* is not equivalent to either *Strong Epistemic Internalism* or the *Hybrid Position*, which are formulated here in terms of access. Indeed, one can endorse *Deontic Mentalism* while rejecting those positions if one thought (e.g.,) that non-accessible mental features play a role in determining the specific deontic status of beliefs. Second, and more importantly, *Deontic Mentalism* merely states that there are no differences in deontic appraisal without differences mentally, but it leaves entirely open *which* mental features are important for states being open to deontic appraisal and *which* mental features determine the particular deontic status of beliefs. But the *Hybrid Position* does not leave these things open. Specifically, it maintains that only some mental features are relevant to the particular deontic status of an agent’s beliefs (i.e. accessible mental features) while other mental features may be relevant to whether the agent’s beliefs have any deontic properties (i.e. any mental feature). So in and of itself *Deontic Mentalism* does not support the *Hybrid Position*.

In any case, my primary purpose here is not to argue that the *Hybrid Position* is false. My primary purpose has been to show that a promising response to Alston’s critique of epistemic deontology comes at a cost: rejecting *Strong Epistemic Internalism*. Proponents of that position may try to retreat to the weaker *Hybrid Position*. But such a retreat will rob them of a major motivation for their views, the *Regulative First-Person Motivation*.

VI. Conclusion

Alston articulated a well-known problem for epistemic deontology composed of two parts: a broadly voluntarist conception of responsibility and strong doxastic involuntarism. One way of developing epistemic deontology in light of Alston’s problem is to reject his strong doxastic involuntarism by embracing a compatibilist account of control, specifically, a reason-responsive account of control. I’ve argued here that this way of developing epistemic deontology in response to Alston’s argument comes as surprising cost: it rules out certain strong internalist views about what epistemic obligations agents have. Of course, this might not suffice to refute such views. But it does show that, for such views to be promising in light of Alston’s problem, they must embrace some other sort of response such as abandoning a voluntarist conception of responsibility. Whether other responses *are* promising is a subject I leave for another time.

Acknowledgements:

For helpful conversation and feedback, I thank Dan Buckley, Jordi Cat, Dave Fisher, Hao Hong, Mark Kaplan, Suzanne Kawamleh, Tufan Kiymaz, Tim Leisz, Adam Leite, Dan Linsenhardt, Sharon Mason, Nick Montgomery, Timothy O’Connor, Luis Oliveira, Harrison Waldo, Phillip Woodward, and three anonymous reviewers.

Bibliography:

Adams, Robert. (1985). “Involuntary Sins.” *Philosophical Review*. 94.1: 3-31.

Alston, William P. (1989). *Epistemic Justification*. Ithaca: Cornell University Press.

- Bayne, Tim. (2016). "Free Will and the Phenomenology of Agency." In the *Routledge Companion to Free Will*, eds. Kevin Timpe, Meghan Griffin, and Neil Levy. Routledge Press.
- Bergmann, Michael. (2006). *Justification without Awareness*. Oxford: Clarendon Press.
- BonJour, Lawrence. (1985). *The Structure of Empirical Knowledge*. Harvard University Press.
- BonJour, Lawrence. (2002). "Internalism and Externalism in Epistemology." In *Oxford Handbook of Epistemology*, ed. P. Moser. Oxford University Press.
- Chisholm, Roderick. (1966). *Theory of Knowledge*. Prentice Hall.
- Chisholm, Roderick. (1980). "A Version of Foundationalism." *Midwest Studies in Philosophy*. 5: 543-64.
- Chrisman, Matthew. (2008). "Ought to Believe." *The Journal of Philosophy*, 105.7: 346-70.
- Christensen, David. (1994). "Conservatism in Epistemology." *Nous*. 28.1: 69-89.
- Cohen, Stewart. (1984). "Justification and Truth." *Philosophical Studies*. 46.3: 279-95.
- Conee, Earl and Richard Feldman. (2001). "Internalism Defended." Reprinted in Conee and Feldman (2004), pp. 53-80.
- Conee, Earl and Richard Feldman. (2004). *Evidentialism*. Oxford: Clarendon Press.
- Conee, Earl and Richard Feldman. (2008). "Evidence." In *Epistemology: New Essays*, ed. Quentin Smith. Oxford University Press.
- Davis, Emmalon. (2016). "Typecasts, Tokens, and Spokespersons." *Hypatia*. 31.3: 485-501.
- Elga, Adam. (2007). "Reflection and Disagreement." *Nous*. 41.3: 478-502.
- Elgin, Catherine. (1996). *Considered Judgment*. Harvard University Press.
- Feldman, Richard. (1988a). "Having Evidence." Reprinted in Conee and Feldman (2004), pp. 219-40.
- Feldman, Richard. (1988b). "Epistemic Obligations." *Philosophical Perspectives*. 2: 235-56.
- Feldman, Richard. (2000). "The Ethics of Belief." Reprinted in Conee and Feldman (2004), pp. 166-196.
- Feldman, Richard. (2002). "Epistemological Duties." In *The Oxford Handbook of Epistemology*, ed. Paul Moser. Oxford: Oxford University Press. 362-84.
- Feldman, Richard. (2006). "Epistemological Puzzles about Disagreement." In *Epistemology Futures*. Ed. Stephen Hetherington. Oxford University Press.

- Fischer, John Martin. (2006). *My Way*. New York: Oxford University Press.
- Fischer, John Martin and Mark Ravizza. (1998). *Responsibility and Control*. Cambridge: Cambridge University Press.
- Foley, Richard. (1993). *Working without a Net*. Oxford University Press.
- Frankfurt, Harry. (1971). "Freedom of the Will and the Concept of a Person." *Journal of Philosophy*. 68.1: 5-20.
- Fricker, Miranda. (2007). *Epistemic Injustice*. Oxford University Press.
- Goldman, Alvin. (1980). "The Internalist Conception of Justification." *Midwest Studies in Philosophy*. 5: 27-52.
- Goldman, Alvin. (1999). *Knowledge in a Social World*. Oxford: Clarendon Press.
- Harman, Gilbert. (1986). *Change in View*. MIT Press.
- Hieronymi, Pamela. (2008). "Responsibility for Believing." *Synthese*. 161.3: 357-73.
- Huemer, Michael. (2007). "Compassionate Phenomenal Conservatism." *Philosophy and Phenomenological Research*. 74.1: 30-55.
- Huemer, Michael. (2013). "Phenomenal Conservatism Uber Alles." In *Seemings and Justification*, ed. Christ Tucker. New York: Oxford University Press, 328-50.
- Kahneman, Daniel. (2011). *Thinking, Fast and Slow*. Farrar, Straus, and Giroux.
- Kelly, Thomas. (2005). "The Epistemic Significance of Disagreement." *Oxford Studies in Epistemology*. 1: 167-96.
- Kornblith, Hilary. (2002). *Knowledge and Its Place in Nature*. Oxford: Clarendon Press.
- Kornblith, Hilary. (2012). *On Reflection*. Oxford: Oxford University Press.
- Kvanvig, Jonathan (1989). "Conservatism and Its Virtues." *Synthese*. 79.1: 143-63.
- Lyons, Jack. (2013). "Should Reliabilists be Worried about Demon Worlds?" *Philosophy and Phenomenological Research*. 86.1: 1-40.
- Madison, B. J. C. (2017). "Epistemic Value and the New Evil Demon." *Pacific Philosophical Quarterly*. 98.1: 89-107.
- McCain, Kevin. (2008). "The Virtues of Epistemic Conservatism." *Synthese*. 164: 185-200.
- McCain, Kevin. (2014). *Evidentialism and Epistemic Justification*. Routledge.
- McGrath, Matthew (2007). "Memory and Epistemic Conservatism." *Synthese*. 157: 1-24.

- Moretti, Luca. (2015). "Phenomenal Conservatism." *Analysis*. 75.2: 296-309.
- Nelson, Mark. (2010). "There are no Positive Epistemic Duties." *Mind*. 119.473: 83-102.
- Parfit, Derek. (2011). *On What Matters*. Oxford University Press.
- Plantinga, Alvin. (1993). *Warrant: The Current Debate*. New York: Oxford University Press.
- Quine, W. V. (1960). *Word and Object*. MIT Press.
- Perrine, Timothy. (under review). "Evidentialism, Knowledge, and Evidence Possession"
- Perrine, Timothy. (2014). "Against Kornblith Against Reflective Knowledge." *Logos & Episteme*. 5(3): 351-60.
- Poston, Ted. (2014). *Reasons and Explanation*. Palgrave.
- Ryan, Sharon. (2003). "Doxastic Compatibilism and the Ethics of Belief." *Philosophical Studies*. 114: 47-79.
- Scanlon, T. M. (1998). *What We Owe Each Other*. Harvard University Press.
- Schmitt, Frederick. (1992). *Knowledge and Belief*. Routledge.
- Schmitt, Meg. (2015). "Freedom and (Theoretical) Reason." *Synthese*. 192.1: 25-41.
- Sklar, Lawrence (1975). "Methodological Conservatism." *Philosophical Review*. 84.3: 374-400.
- Smith, Angela. (2005). "Responsibility for Attitudes." *Ethics*. 115.2: 236-71.
- Silins, Nicholas. (2005). "Deception and Evidence." *Philosophical Perspectives*. 19: 375-405.
- Steup, Matthias. (2000). "Doxastic Voluntarism and Epistemic Deontology." *Acta Analytica*. 15.24: 25-56.
- Steup, Matthias. (2008). "Doxastic Freedom." *Synthese*. 161.3: 375-392.
- Steup, Matthias. (2011). "Belief, Voluntariness, and Intentionality." *Dialectica*. 65.4: 537-59.
- Steup, Matthias. (2012). "Belief Control and Intentionality." *Synthese*. 188.2: 145-63.
- Steup, Matthias. (2017). "Believing Intentionally." *Synthese*. 194.8: 2673-2694.
- Tucker, Chris. (2010). "Why Open-Minded People Should Endorse Dogmatism." *Philosophical Perspectives*. 24: 529-45.
- Vahid, Hamid (2004). "Varieties of Epistemic Conservatism." *Synthese*. 141.1: 97-122.

Weatherston, Brian. (2008). "Deontology and Descartes's Demon." *The Journal of Philosophy*. 105.9: 540-69.

Wolterstorff, Nicholas. (1996). *John Locke and the Ethics of Belief*. Cambridge University Press.