

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.

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.

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.

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 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 Alston (1988) has called “the deontological conception of epistemic justification,” on which beliefs are aptly evaluated using deontic concepts like *ought*, *obligation*, *permission*, and the like.

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 (1974), Stalnaker (1987), and Dennett (1987), among others.² Here is Stalnaker (p. 4) 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.

² 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.

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 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.³)

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 Christensen (2001) and Meacham and Weisberg (2011) for related discussion). We

³ See Hájek and Smithson (2012) for discussion. They also highlight the parallel between Lewis' theory of mental states and his theory of laws and chances.

also want the agent's beliefs to come out rational in light of her evidence⁴, 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.⁵ 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,⁶ how important it

⁴ 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.

⁵ 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.

⁶ 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 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

is that the beliefs 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. Lewis (1974, 337) 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*⁷ 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**.

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

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.

⁷ 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.

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.⁸

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 (1994, 478) 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.

⁸ 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 (Holton (2014), Ross and Schroeder (2014)). 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.

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

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, as 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 (1994, 480):

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.

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 (1994, 482) 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.”

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” (ibid). 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” (ibid).

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 (1980). 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).⁹ 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 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

⁹ 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.

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 (Feldman (2000) 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 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.¹⁰

¹⁰ Greco (2012) 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 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” (p. 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 Millikan's (1989) biosemantic picture of belief), 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.

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,¹¹ 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 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.¹² 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. Griffiths (1997) argues that certain 'core' emotions, such as fear, surprise, anger, disgust,

¹¹ Recall that in the quote from Stalnaker above, he mentions 'idle wishes, passive hopes, and theoretical belief' in addition to ordinary beliefs.

¹² See Kim (2007) for discussion.

sadness, and joy, are automatic responses to environmental stimuli characterized by their associated physiological changes and their evolutionary phylogeny (Darwin (1872) found 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.

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. Alston (1988) 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 Peels (2014) 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).¹³ 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 voluntary

¹³ Alston 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" (1988, 259).

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.¹⁴

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. Lewis (1981) 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 L specifying the (deterministic) laws of nature, and H and L 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.

¹⁴ Setiya (2008) 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.

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 L 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, L 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 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 is 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.)¹⁵

¹⁵ 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.

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 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’ (1988, 277). Here is how he puts it (p. 278):

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.

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.

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; see Ryan (2003), McHugh (2014), and Steup (2000).) 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.¹⁶ 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 .¹⁷

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

¹⁶ 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.

¹⁷ 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 (1988, 262) have thought that you needn't *always* have freedom or voluntary control over your beliefs in order for deontic concepts to be applicable to beliefs; it is enough that you sometimes, or at least often, have such freedom or voluntary control.

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.

References

- William Alston. "The Deontological Conception of Epistemic Justification." *Philosophical Perspectives* 2 [1988]: 257-99.
- David Christensen. "Preference-Based Arguments for Probabilism." *Philosophy of Science* 68 [2001]: 356-76.
- Charles Darwin. *The Expression of the Emotions in Man and Animals*. Originally published 1872. Reprinted by University of Chicago Press, 1965.
- Daniel Dennett. *The Intentional Stance*. Cambridge: MIT Press, 1987.
- Richard Feldman. "The Ethics of Belief." *Philosophy and Phenomenological Research* 60 [2000]: 667-95.
- Daniel Greco. "The Impossibility of Skepticism." *Philosophical Review* 121 [2012]: 317-58.
- Paul Griffiths. *What Emotions Really Are*. Chicago: University of Chicago Press, 1997.
- Alan Hájek and Michael Smithson. "Rationality and Indeterminate Probabilities." *Synthese* 187 [2012]: 33-48.
- Richard Holton. "Intention as a Model for Belief." In *Rational and Social Agency: Essays on the Philosophy of Michael Bratman*, ed. Vargas and Yaffe (Oxford University Press, 2014).
- Sharon Ryan. "Doxastic Compatibilism and the Ethics of Belief." *Philosophical Studies* 114 [2003]: 47-79.
- Jaegwon Kim. *Physicalism, or Something Near Enough*. Princeton: Princeton University Press, 2007.
- David Lewis. "Radical Interpretation." *Synthese* 27 [1974]: 331-44.
- David Lewis. "A Subjectivist's Guide to Objective Chance." In *Studies in Inductive Logic and Probability*, ed Richard Jeffrey (University of California Press, 1980), 83-132.
- David Lewis. "Are We Free to Break the Laws?" *Theoria* 47 [1981]: 113-21.
- David Lewis. "Humean Supervenience Debugged." *Mind* 103 [1994]: 473-90.
- Conor McHugh. "Exercising Doxastic Freedom." *Philosophy and Phenomenological Research* 88 [2014]: 1-37.
- Christopher Meacham and Jonathan Weisberg. "Representation Theorems and the Foundations of Decision Theory." *Australasian Journal of Philosophy* 89 [2011]: 641-63.

Rik Peels. "Believing at Will is Possible." *Australasian Journal of Philosophy* 88 [2014]: 1-18.

Jacob Ross and Mark Schroeder. "Belief, Credence, and Pragmatic Encroachment." *Philosophy and Phenomenological Research* 88 [2014]: 259-88.

Kieran Setiya. "Believing at Will." *Midwest Studies in Philosophy* 32 [2008]: 36-52.

Robert Stalnaker. *Inquiry*. Cambridge: MIT Press, 1987.

Matthias Steup. "Doxastic Voluntarism and Epistemic Deontology." *Acta Analytica* 15 [2000]: 25-56.