Believing at Will

Pamela Hieronymi
hieronymi@ucla.edu
February 21, 2011

It has seemed to many philosophers—perhaps to most—that believing is not voluntary, that we cannot believe at will. It has seemed to many of these that this inability is not a merely contingent psychological limitation but rather is a deep fact about belief, perhaps a conceptual limitation. But it has been very difficult to say exactly why we cannot believe at will.

I earlier offered an account of why we cannot believe at will.¹ I argued that nothing could qualify both as having been done “at will,” in the relevant sense, and as a belief. Thus, no believer could believe at will. If my arguments are correct, our inability to believe at will reveals no genuine lack in our powers of mind, any more than an inability to draw a square circle reveals a lack of artistic skill.

My account has been recently criticized by Kieran Setiya, who has provided an account of his own.² Here I revisit and defend my account, hopefully in a way that will both make my thought clearer and illumine some of the broader differences between Setiya’s approach and my own. I then briefly consider Setiya’s own argument, in part to further develop the contrast.³

THE PROBLEM: WHY CAN’T WE BELIEVE AT WILL?

To begin, we should examine why it has been so difficult to explain why we cannot believe at will. For this, I find an old exchange between Bernard Williams and Jonathan Bennett helpful. Both Williams and Bennett agree that believing is involuntary, and, furthermore, that its involuntariness is,
as Williams puts it, “not a contingent fact.” Williams illustrates this by contrasting believing with blushing, thinking our inability to blush at will is a merely contingent fact:

> We must agree that there are cases of what we may call contingent limitations on the will. For instance, suppose someone says that he cannot blush at will. What would it be to blush at will? You could put yourself in a situation which you would guess would make you blush. That is getting yourself to blush—by a route—but it could not possibly count as blushing at will. Consider next the man who brings it about that he blushes by thinking of an embarrassing scene. That is getting a bit nearer to blushing at will, but is perhaps best described as making oneself blush at will. The best candidate of all would be someone who could blush in much the way that one can hold one’s breath. I do not know whether people can do that, but if they cannot, it will be a contingent fact that they cannot.⁴

Although I think the contrast with blushing is misleading in several different and important ways, it does highlight an important point: bringing yourself to blush “by a route,” as Williams puts it, “count not possibly count as blushing at will.” This point is important because you can also bring yourself to believe, by a route. You might put yourself in front of compelling evidence, change the world to make the belief obviously true, induce in yourself amnesia about the relevant bit of history, take some medication to quell your fears, or perhaps even take some science-fictional belief-inducing pill. But believing in any of these ways could not possibly count as believing at will. It would be, instead, bringing yourself to believe or making yourself believe. Bringing yourself to believe, even to a specific belief, can be done—though, like most everything that we do, the circumstances must be right, it is often difficult, and we do not always succeed. Thus, although neither believing nor blushing are done at will, both are subject to one’s will, insofar as one can take action to elicit or to extinguish either one. On this everyone agrees. But being subject to the will in this way—being manageable—is not the same as being voluntary. On this, everyone also agrees. Part of the difficulty of explaining why we cannot believe voluntarily, or at will, then, lies in understanding what would count as believing voluntarily, or at will.

---

We just noted that, if one is to take action and successfully bring oneself to a specific belief (by, say, uncovering evidence or inducing amnesia or changing the world), the circumstances must be especially cooperative. This might tempt us to think that our supposed inability to believe at will can be accounted for simply by appealing to the difficulty and delicacy of bringing oneself to a specific belief. Perhaps, when we say that a person can do something “at will,” we mean that he can do it easily or readily, in most circumstances. Raising one’s right hand is something that, for most people, can be done easily, in most circumstances. But notice, were this what we meant by “at will,” then not only could we not believe at will, but we also could not signal for a left hand turn, put on our favorite shoes, or enjoy a cup of coffee at will—because we cannot do these easily, in most circumstances. We must be at an intersection, near our shoes, or around coffee and in a good enough mood. But whatever it is that divides believing from raising your right hand also, it seems, divides believing from putting on your favorite shoes or enjoying a cup of coffee. The ability to a thing “at will,” at issue, is not the ability to do a thing easily, in most circumstances.

One might be tempted, next, to think that the relevant ability is the ability to do a thing easily or readily, in favorable or typical circumstances. I can easily put on my shoes or enjoy a cup of coffee, in favorable or typical circumstances, but, one might think, I cannot easily bring myself to believe, even in favorable or typical circumstances. Bringing yourself to believe may thus seem to be a bit like hitting a bullseye: difficult or uncertain, even in favorable or typical circumstances. And it seems right to say that most people cannot hit a bullseye at will, due to the difficulty and uncertainty of the task. So one might be tempted to think that we cannot believe “at will,” due to the difficulty or delicacy of the task.

But I hope this temptation will also quickly pass: Notice, first, that bringing yourself to a specific belief need not always be difficult or delicate. As pointed out by Feldman, if you want to believe the lights are on in your office, you can throw the switch. As pointed out by Kelly, if you

---

want to believe your children are safe at home, you can make a phone call.\textsuperscript{6} If all we meant, in saying that you cannot believe at will, is that doing so is difficult or uncertain, then we should say that, in these easy cases, you \textit{can} believe at will. But this seems wrong. So, I hope it plain that what divides believing from raising one's right hand is not the same as that which divides raising one's right hand from hitting a bullseye. The limitation seems, to use Williams' language, deeper.

In our attempt to locate the relevant sense of “at will,” one might next look for help in the phrase “fire at will,” said to soldiers in battle. As I understand it, this command contrasts with firing all together, in a volley—firing only when commanded. So it means something like, “fire when you see fit,” or “fire when you judge it called for.” But, of course, in this sense, we can and do believe at will.\textsuperscript{7}

So, just what it is to do something “at will,” in the relevant sense, remains obscure. Williams does not pursue this question; he does not ask (as he did with blushing), “What would it be to believe at will?” Instead, he seems to assume that we would either blush or believe at will if we could do either one “in much the way that one can hold one’s breath.” He seems to think that our inability to blush at will is a merely contingent limitation of something like our wiring—just as some people can wiggle their ears without using their hands while others cannot, and this is a merely contingent matter, so some people might be able to blush at will, though most of us cannot, and this is a merely contingent matter.\textsuperscript{8} In contrast, he thinks our inability to believe at will is not a merely contingent fact, not a merely contingent feature of our wiring or our psychology.


\textsuperscript{7} One might think the point is that you cannot believe \textit{only} when you see fit or judge it called for: you have to believe whatever is obviously so, whether you want to believe it or not. This is, I think, getting near the point. But notice, you also cannot fire your weapon only when you see fit for judge it called for: you cannot rule out firing inadvertently or against your better judgment. This does not impugn the claim that you can \textit{also} fire at will. Yet, when we believe because we judge it called for, we are not, it seems, believing at will.

\textsuperscript{8} I believe Williams is mistaken about blushing. See note 37.
Williams supports his conviction with the following bit of argument:

Belief cannot be like that; it is not a contingent fact that I cannot bring it about, just like that, that I believe something, as it is a contingent fact that I cannot bring it about, just like that, that I’m blushing. Why is this? One reason is connected with the characteristic of beliefs that they aim at truth. If I could acquire a belief at will, I could acquire it whether it was true or not. If in full consciousness I could will to acquire a ‘belief’ irrespective of its truth, it is unclear that before the event I could seriously think of it as a belief, i.e., as something purporting to represent reality. At the very least, there must be a restriction on what is the case after the event; since I could not then, in full consciousness, regard this as a belief of mine, i.e., as something I take to be true, and also know that I acquired it at will. With regard to no belief could I know . . . that I had acquired it at will. But if I can acquire beliefs at will, I must know I am able to do this; and could I know that I was capable of this feat, if with regard to every feat of this kind I had performed I necessarily had to believe that it had not taken place?

Roughly, the argument is this: because beliefs purport to represent reality, any exercise of an ability to believe at will would have to be self-effacing; it would have to include something like a brief bout of amnesia. But, if each exercise of this ability is self-effacing, one would not know that one had the ability. And if one does not know one is able to believe at will, then one is not able to do so.

By pointing to the fact that beliefs purport to represent reality, Williams seems to have put his finger on something crucial, something that promises to explain our inability to believe at will. And yet his argument does not succeed. Jonathan Bennett (who shares Williams’ the conviction that no believer could believe at will) put forward a counter-example to Williams’ argument (and confessed his inability to provide a better argument to the same conclusion):

Credam is a community whose members can [decide to believe]. It doesn’t happen often, because they don’t often think: “I don’t believe that p, but it would be good if I did.” Still, such thoughts come to them occasionally, and on some of those occasions the person succumbs to temptation and wills himself to have the desired belief. . . . When a Credamite gets a belief in this way, he forgets that this is how he came by it. The belief is always one that he has entertained and has thought to have some evidence in its favor; though in the past he has rated the counter-evidence more highly, he could sanely have inclined the other way. . . . The trick cannot be worked if the protective forgetfulness would require that the rest of the person’s beliefs be drastically rearranged. . . . After successfully willing himself to have a certain belief, a Credamite may later get evidence that that is what he has done; e.g., someone may tell him. Then he either rejects the evidence or loses the belief. . . . So

9 Williams, "Deciding to Believe," 148.
each Credamite knows that he sometimes wills himself to believe something, even though it is never true that he now has a belief which he now remembers having willed himself to acquire.  

By appealing to a community, Bennett shows that the necessarily self-effacing nature of the supposed ability does not show it impossible to exercise. If purporting to represent reality rules out voluntariness, we are still in the dark about why. And still in the dark, as well, about the relevant notion of voluntariness.

INTRODUCING AN ALTERNATIVE

I hope to succeed where Williams and Bennett did not; I think I can say why our inability to believe at will is not a merely contingent fact of our psychology. In doing so, I will locate the relevant notion of voluntariness.

However, before giving my argument, I suspect it will help to introduce the account by previewing the interpretation of Credamites it will yield. Despite thinking that no believer can believe voluntarily, I will not try to argue that what Bennett imagines is impossible. Rather, I will suggest that what Bennett imagines is not a case of believing voluntarily. What the Credamites do, instead, is to voluntarily bring it about that they believe. That is to say, their belief is not, itself, voluntary, but rather is the product of the voluntary action of making themselves believe. While we ordinary mortals can make ourselves believe through various complicated, clunky processes—by seeking out evidence or taking a pill or changing the world to make the belief obviously true—the Credamites are able to do it just by deciding to do it, as a so-called “basic” action. We do it by a route; they do it directly. The Credamites are thus simply the limiting case of making yourself believe, as the route shortens and finally disappears. But, I will now suggest, making yourself believe, even as a basic action, is not believing at will, in the relevant sense.

Williams’ examples have misled us. He suggests that someone who could make her face flush in the same way that most of us can hold our breath would have the ability to blush at will, and he thereby tempts us to confuse the question of whether an activity is voluntary with the question of whether an action is, as it is sometimes put, “basic”—whether an action is one that we can do, in Williams’ words, “just like that,” without having to undertake other, prior, intermediary, or component actions.\(^{11}\) It is the first question that should interest us, so I will here distinguish it from the second.

To start: what is a basic action? Though there have been doubts raised, lately, about whether there are any basic actions, or whether that notion is sound,\(^{12}\) a rough sense of it will do, for our purposes: roughly, an action is basic if it can be done without first having to do something else. So, holding your breath and raising your right hand are, for most people, “basic” actions. Most people do not need to use their left hand in order to raise their right, and whatever muscular movements go into holding one’s breathing do not stand out, themselves, as separate actions. In contrast, making soup, rearranging the furniture, and traveling by air are non-basic actions.

When someone can do something unusual as a basic action, it does seem natural to say that person can do that thing “at will.” Some people—my father was one—can wiggle their ears as a basic action. I cannot. It seems natural to say that my father could wiggle his ears “at will,” while I cannot. But notice that what one means, when one says that I cannot wiggle my ears “at will,” is simply that, if I want my ears to wiggle, I will have to use my hands. So long as we do not rule out manual ear wiggling, I certainly can wiggle my ears at will. (In fact, I can do so easily, in almost any circumstance.) What I cannot do (at all) is that thing my father could do, namely, wiggle my ears without using my hands, or wiggle my ears as a basic action. And this, it seems, is a contingent feature of my physiology.

\(^{11}\) It may be worth noting that you can make your face flush by holding your breath long enough. Blushing, one might think is another matter. Cf. note 37.

\(^{12}\) Doug Lavin has raised such doubts in a lecture titled “Must There Be Basic Action?”
But when we say that we cannot believe at will, we do not simply mean that we cannot believe as a basic action. We mean to say that believing is not like acting, at all—not like any voluntary action, whether basic or not. That is, I hope it clear and uncontroversial enough that both basic and non-basic actions are voluntary in whatever sense believing is not: whatever divides believing from raising one’s right hand or my father’s trick also divides believing from manual ear-wiggling, making soup, rearranging the furniture, or traveling by air. Were one endowed with the God-like ability to make soup “just like that,” by deciding to do so, without all the troublesome intervening labor—if one could simply decide, “let there be soup!” and have it be so—one’s soup-making would thereby have been made much easier. But it would not thereby have been made any more voluntary.\(^{13}\) The complexity of the activity is not the issue.

So we must distinguish the question of whether an activity is or could be voluntary, in the sense at issue, from the question of whether an action is or could be basic. When we ask why we cannot believe at will, we are not simply asking whether the voluntary, non-basic action of bringing yourself to believe could become a basic one. (If we were simply interested in whether we could bring ourselves to believe as a basic action, “just like that,” then Bennett’s example makes a compelling case that believers with certain fancy psychological equipment could do so.\(^{14}\) We are, rather, asking why believing is not like either raising your right hand or making soup—why it is not voluntary in a sense that both hand-raising and soup-making are. Believing does not seem to be voluntary at all. (Other activities that do not seem to be voluntary include loving, hating, trusting, resenting, feeling blue, and seeing red.)

\(^{13}\) Perhaps one wants to say that God can make soup at will, while I cannot. One would then be making the point made above, in saying my father can wiggle his ears at will while I cannot: God can create soup as a basic action, while, for me, it is a bit more difficult.

\(^{14}\) Of course, it must be granted that, because beliefs “purport to represent reality,” the circumstances in which this fancy ability could be exercised would have to be limited in the ways Bennett outlines (the belief would have to be broadly plausible, the ability would have to come paired with very short-term amnesia, etc.). But limitations of various sorts attend all our actions, without licensing the claim that they cannot be done voluntarily or even as a basic action.
Once we distinguish the question of whether an action is basic from the question of whether an activity is voluntary, the Credamites no longer present a clear counter-examples to the claim that you cannot believe at will; we can instead interpret them as bringing it about that they believe, as a basic action.

What would it be, then, to believe voluntarily, or at will? I suggest that it would be to believe in the way we ordinarily act—for believing not to be the product of an action of belief-making, but rather to be an intentional action in its own right.\(^{15}\) However, I will argue, even though you may successfully execute intentions regarding your beliefs, with the result that you gain the belief you intended to gain, and even though (if you are a Credamite) you may be able to execute such an intention with godlike ease, you cannot believe as an intentional action in its own right.\(^{16}\) This is because nothing could qualify as both an intentional action and a belief. (So, you cannot, properly speaking, form and execute an intention to believe. You can, at best, form and execute an intention to bring it about that you believe.) Thus, the closest we could approximate believing at will would be to bring ourselves to believe as a basic action. But everyone should agree both that bringing yourself to believe is possible and that it is not believing at will.

**WHY ISN’T BELIEVING AN ACTION?**

So, why could believing not be an action in its own right? Why could nothing quality as both an intentional action and a believing? Briefly, because believing brings with it a different sort of commitment, or answerability, than does acting. This requires elaboration.

When you believe, you are committed to the truth of your belief. That is, if you believe that the butler did it, you are vulnerable to questions and criticisms that would be satisfied by reasons that

\(^{15}\) Williams’ language does not suggest this interpretation of believing at will. He uses with the noun, treating a belief as an item that one might acquire or produce, not as an activity that might be itself be voluntary.

\(^{16}\) Nor, to anticipate Setiya’s criticisms, can you form a belief as an action in its own right. The reader can substitute “form the belief” for “believe” as she or he sees fit. I return to this below.
bear (or that you take to bear, or that you ought to take to bear, or that it would be reasonable to take to bear) on the butler's guilt.\textsuperscript{17} I will express this by saying that, if you believe \( p \), you are committed to a positive answer to the question of whether \( p \).

In contrast, whenever you act intentionally, or whenever you intend to act, you are committed to so acting. If you run for office intentionally, or if you intend to run for office, you are vulnerable to questions and criticisms that would be satisfied by reasons that bear (or that you take to bear, or that you ought to take to bear, or that it would be reasonable to take to bear) on the question of whether to run. Again, I think this can be clearly put by saying that, if you \( x \) intentionally, or intend to \( x \), then you are committed to a positive answer to the question of whether to \( x \).\textsuperscript{18}

Notice, further, that when one decides to act, one is committed to one's action—including all parts of a complex action—from the point of decision. So, if you decide to make soup, you are therein committed to the entire, complex activity of soup-making: to gathering ingredients, turning on the burner, dirtying a pot, etc. Even if you never get as far as turning on the stove, you are still, simply in \textit{intending} to make soup, vulnerable to questions about whether dirtying a pot is worth doing. Likewise, were you to suffer some sudden, unexpected paralysis just as you attempt to raise your right hand, such that your hand never moves, you are still committed to raising your right hand, simply in intending to do so. In general, if one intends to \( x \), one is therein committed to \( x \)-ing, including all of \( x \)-ing's intentional parts. (I take this to be an uncontroversial feature of action, as we actually understand it.\textsuperscript{19} I will, below, suggest that this feature accounts for our sense that action is voluntary, in the sense presently at issue.)

\textsuperscript{17} The parenthetical allows for the wide range of kinds of questions and criticisms to which one is subject, the variety of which is not important for present purposes. You can be asked for your reasons for believing. You can be criticized if your reasons are not reasonable. Your believing can evaluated in a different way if you reasons are untrue, or if they do not in fact bear on the butler's guilt, even if it was reasonable so to take them. Your belief might be evaluated in yet a different way by appeal to reasons that show it likely or unlikely, even if you do not have access to those reasons.

\textsuperscript{18} I am, here, ignoring additional complexity about, e.g., foreseen side effects. Cf. note 20, below.

\textsuperscript{19} As we do understand it, not as we must. See note 28.
Moreover, actions often have consequences or products that are not, themselves, part of the action. Making soup produces some soup, but the soup is not itself part of the action of soup-making. The soup is its product. Traveling by air creates some nasty emissions, but those emissions are not, themselves, part of the action of traveling by air—though, arguably, generating those emissions is part of that action.\(^{20}\) Since neither soup nor emissions are themselves actions (indeed, neither is even an activity), you are not, in deciding to make soup or travel by air, committing to doing them. Rather, insofar as such consequences are either intended or foreseeable (or, perhaps, foreseeable), you are committed, from the point of decision, to bringing them about, as a product of the action you have decided to do.\(^{21}\)

So, in believing \(\phi\), one is committed to a positive answer to the question of whether \(\phi\), while, in either \(x\)-ing intentionally or intending to \(x\), one is committed to a positive answer to the question of whether to \(x\). Why should this difference in commitment preclude believing as an action in its own right? Suppose we grant it. We thereby grant that believing \(\phi\) as an action in its own right will require committing yourself not only to a positive answer to the question of whether to believe \(\phi\), but also to a positive answer to the question of whether \(\phi\). Why is that a problem?

Notice, first, that in asking “whether to believe \(\phi\),” you could be entertaining two starkly different questions. You might be thinking simply about whether \(\phi\). On this first interpretation, the question

---

\(^{20}\) It is often said that creating the emissions is not something you intended, because, if there were a way to travel without the emissions, you would prefer it (or, at least, be indifferent to it). Thus, not everything done intentionally is intended (you intentionally create the emissions, though you did not intend to do so). Throughout this essay I allow the notion of commitment to be interpreted more broadly or narrowly (hence the parenthetical remarks about whether the reasons are ones that you take to bear, or that you ought to take to bear, etc.). Since I am trying to show that one cannot incur a commitment to the truth of \(\phi\) as a part of an action, I should interpret the commitments involved in action as broadly as possible (to make my task harder). Thus I will understand one to be committed to everything one does intentionally. (For important discussion of these issues, see Michael E. Bratman, Intention, Plans, and Practical Reason (Cambridge: Cambridge University Press, 1987). I am grateful to both Michael Bratman and Grant Rozeboom for pressing me on related issues.

\(^{21}\) Setiya considers the difference between those things that are constitutive of an action and those that are the products or results of an action. He suggests, in note 7, that the difference turns on the “nature and unity of events as such.” I strongly suspect that the difference will instead be determined by the extent to which we think it apt to hold someone answerable.
of whether to believe \( p \) is just the ordinary, “epistemic” question; it will be settled for reasons that you take to bear on whether \( p \). If you think about and settle the question (positively) under this interpretation, you will believe \( p \) in an ordinary way, and will incur only the usual commitment to the truth of \( p \). This is not believing at will.

To believe at will, or as an action in its own right, you must incur a commitment to a positive answer to the question of whether to believe \( p \) on a second, practical interpretation: you must consider and settle the question of whether to believe \( p \) in the same way that you consider and settle the question of whether to run for office or raise your right hand. The reasons that you take to bear on the question of whether to believe, under this interpretation, will be reasons that you take to show, not that \( p \), but rather that believing \( p \) is somehow worth doing. These are sometimes called pragmatic reasons. I have called them *extrinsic reasons* for the belief that \( p \). They are, again, reasons that you do not take to be sufficient to settle the question of whether \( p \), but which you nonetheless take to settle the question of whether believing \( p \) would be a good idea, reasons you take to show believing \( p \) in some way good, useful, valuable, important, satisfying, or some such.\(^{22}\) (These are the same reasons you would take to bear on the question of whether to bring it about that you believe \( p \), were you left with only that more prosaic method of belief acquisition.) I will gloss the question of whether to believe \( p \), so interpreted, as the question of whether believing \( p \) is somehow worth

\(^{22}\) Two points are worth noting. First, you may take the overall likelihood or probably of \( p \) into consideration, when asking whether believing \( p \) would be, all things considered, good, useful, etc. You may, that is, aver to some “epistemic” reasons. In this case, though, you are taking considerations which you take to bear on but not settle the question of whether \( p \) to bear, in turn, on whether believing \( p \) is, all things considered, worth doing. Thus they are still extrinsic reasons. Second, in certain unusual cases, the question of whether believing \( p \) is worth doing may come to bear on, and help to settle, the question of whether \( p \) (most starkly, for the case in which \( p \) is “this belief is worth having”). In that case, one may settle the question of whether \( p \) for reasons that one takes to show believing \( p \) is worth doing. In this case, these are *not* extrinsic reasons. (I discuss these points, and give a less stark illustration of the second, in Hieronymi, "Controlling Attitudes."; Pamela Hieronymi, "The Wrong Kind of Reason," *The Journal of Philosophy* 102, no. 9 (2005).) The general lesson is that whether a reason is extrinsic or not depends, not on its content, so to speak, but rather on the question it is taken to settle, and that the question of whether to believe \( p \), understood practically, cannot be the question of whether \( p \).
If you are to believe at will, you must believe for reasons that you take to settle this question, i.e., extrinsic reasons.

So, if you are to believe at will, as an action in its own right, you must be committed to both a positive answer to the question of whether to believe \( p \), under a practical interpretation, and to a positive answer to the question of whether \( p \). But again, why is that a problem? Why could you not commit to worthiness of believing \( p \) and to the truth of \( p \), in a single action (perhaps an action that unfolds in stages)?

The problem is not that the two commitments are incompatible. (You could readily both believe that believing \( p \) is worth doing and believe that \( p \) is true.) Nor is the problem that no state of mind or activity could involve more than one kind of commitment. Certain states of mind and activities (such as, e.g., resentment or trust) do involve multiple commitments. Rather, the problem is that, to count as believing at will, it is not enough simply to have both commitments. The mere co-occurrence of these commitments will not amount to believing at will. Rather, to believe at will, you

---

23 Nothing should turn on the particular choice of “somehow worth doing.” It is my attempt to render the question of whether to believe \( p \) practical. This attempt is fraught with difficulty because, as I will explain below, I actually take the correct formulation of the practical question about \( x \)-ing to be not whether \( x \)-ing is “somehow worth doing,” but simply whether to \( x \). Moreover, I am in the process of arguing that believing is not a legitimate value for \( x \), in this question. Thus the proper way to render the question of whether to believe \( p \) practical would be, not whether believing is somehow worth doing, but rather whether to bring it about that one believes. However, making these claims at this point would be question-begging. Thus, I need to formulate, in some way, a practical question about the purported action of believing and distinguish it from the ordinary epistemic question. I formulate it as “whether believing is somehow worth doing.” The reader can reformulate as he or she sees fit (whether believing is desirable, appropriate, to be done, useful, important, etc.).

One might think there is room for something like an intermediate interpretation of the question about whether to believe \( p \), where one asks explicitly about one’s own belief (not merely about \( p \)), but somehow takes oneself to be constrained, in arriving at an answer, to reasons that one takes to be reasons that one takes to be on whether \( p \). (Such a view is advocated by Nishi Shah and J. David Velleman, “Doxastic Deliberation,” Philosophical Review 114, no. 4 (2005).) I would not allow for such an intermediate interpretation: it seems to me that, by taking yourself to be constrained your the reasons, you show yourself to be (in fact) addressing the simpler question, about whether \( p \). But this disagreement can be overlooked for present purposes: even those who allow for the intermediate interpretation would not think it is a practical interpretation—answering it is not believing at will.

24 I am not overlooking the possibility that you might believe for no reason, or for no particular reason. However, it would be odd to think that believing for no reason amounts to believing at will (or, that you could believe at will for no reason). Believing for no reason is, rather, finding yourself with a belief. To be in a position to believe at will, you need to take some reason, however slight, to count in favor of believing \( p \). (Bennett agrees that to believe at will is to believe for practical reasons.)
must believe *because* you decided to believe. And so you must incur the commitment to \( p \)'s truth

*because* you settled the practical question.

One straightforward way of achieving the required consequence relation would be for the commitment to the truth of \( p \) to be result or effect of the practical commitment: you could make the practical commitment, which, in turn, could somehow cause or bring about the commitment to \( p \)'s truth. This is a natural way to understand what it would be to control your beliefs by an exercise of will. But, remember, you commit to all parts of an action from the point of decision. Thus, if the commitment to \( p \)'s truth is merely the result or effect of the commitment to the practical question, believing \( p \) will be understood, not as a part of the action you have decided upon, but rather as the product or consequence of that action—an action best described as bringing it about that you believe or making yourself believe. So, on this natural way of understanding what it would be to control your beliefs, they are subject to your will—manageable—but not themselves voluntary. They stand to your decision to believe as soup stands to your decision to make soup.

To illustrate, imagine that something goes wrong with a Credamite’s special ability: after deciding to believe that \( p \), expected result does not follow. The Credamite suffers the equivalent of unexpected paralysis and doesn’t believe \( p \). In such a case, it seems, the Credamite is not vulnerable to the questions and criticisms to which one is vulnerable when one believes \( p \). While you could rightly ask him why he thought it a good idea to believe \( p \), or why he meant to believe \( p \)—that is, you could rightly ask him for reasons that would bear on whether to bring it about that he believes \( p \)—you could not rightly ask him why he thinks \( p \) is true. After all, he doesn’t yet think \( p \) is true. He has, thus far, only been convinced that it would be good to believe \( p \). He is in the same position you would be if, having decided to call your children to assure yourself that they are safely home, you find that your phone does not work. You are not yet committed to the truth of the claim that your children are home. You are only answerable for reasons that bear on whether to (try to) give yourself that belief—your reasons for placing the call. Both you and this Credamite will *become*
answerable for reasons bearing on the truth of $p$ only if you successfully execute your intention and believe. But, again, you incur a commitment to all parts of an action, simply in intending to act, regardless of whether your action is successful. Thus, it seems, the target belief is best understood, not as a part of the action you intended, an action of believing (as an action in its own right), but rather as an intended consequence of a different action, the product of the action of bringing it about that you believe.

Could we achieve the required consequence relation in a different way? Could we provide an interpretation in which the Credamite’s believing is a consequence of his decision to believe (as it must be, if he is to believe at will), and yet is not simply the product of the action decided upon, but rather is (part of) the action itself (as it must be, if believing is to be an action in its own right)? It seems that, to arrive at such an interpretation, we would have to understand the Credamite as committing to the truth of $p$ (and not just to bringing about such a commitment) by deciding or in his decision to believe. He would have to settle the question of whether $p$ by or in settling the question of whether believing $p$ is somehow worth doing. But the Credamite, himself, does not take the reasons at hand to settle the question of whether $p$ (if he did, he would believe in the ordinary way). So, in order to believe at will, as an action in its own right, the Credamite would have to settle the question of whether $p$ by finding convincing reasons that he does not, himself, take to settle that question. But this he cannot do.

Stepping away from the example: If you are to believe $p$ at will, you must become committed to the truth of $p$ because you decided to believe $p$. And yet, to believe $p$ as an action in its own right, the commitment to the truth of $p$ cannot be merely the result or product of deciding to believe—it

---

25 If he were somehow to do this, the intention formed would not need to be executed: he would already believe. I take this to be another indication that beliefs are not proper objects of intention.

26 A Credamite might be, so to speak, akratic in his lack of belief: he might believe that the available reasons show that $p$ without actually taking them to settle the question of whether $p$ and so believing $p$. In such a case, he will be in a position to believe at will only if he also takes there to be some reason to correct himself in this matter. Thus, even in such a case, he will settle the practical question and decide to believe for extrinsic reasons—reasons that he does not take to settle the question of whether $p$ but nonetheless takes to show that believing $p$ is somehow worth doing.
cannot stand to your decision to believe as soup stands to a decision to make soup. Thus, in order
to believe $p$ as an action in its own right, you would have to commit yourself to the truth of $p$ in
deciding to believe—in answering the practical question. But you do not yourself, take the reasons
at hand to show that $p$ (if you did, you would therein believe in the ordinary way). So, in order to
believe as an action in its own right, you would have to settle the question of whether to believe $p$
for reasons that you do not, yourself, take to settle the question of whether $p$, and therein commit
yourself to a positive answer to the question of whether $p$. But this you cannot do. You cannot
become committed to an answer to a question by finding convincing reasons that you, yourself, do
not take to settle that question. Rather, by finding certain reasons convincing, by taking them to
settle some question, you therein become committed to an answer to whichever question you took
them to settle. So, by finding convincing reasons that you take to show that believing $p$ is worth
doing, you become committed to a positive answer to the question of whether believing $p$ is worth
doing—you have settled a question about the desirability of a certain state of mind. You may now
believe that believing $p$ is worth doing, or you may form an intention regarding the belief. But you
will not, by finding convincing reasons that you take only to show that believing $p$ is worth doing,
therein become committed to the truth of $p$.27 Yet this is what you would have to do, to believe at
will.

---

27 One might try to ask, Why can’t you commit yourself to the truth of $p$ by finding convincing reasons that you do not
take to show it true? But I do not know how to answer this question, because it seems to me not to make good sense.
It seems to me a bit like asking, why can’t I subtract two by adding one? You cannot, but this is not because of any
obstacle or due to any shortfall in your abilities to which I could point, in answer to your question. It is just that adding
one will never be subtracting two. Likewise, to be committed to the truth of $p$ is to be open to the questions and
criticisms associated with believing $p$—to those questions and criticisms that would be satisfied by reasons that bear (or
that one might take to bear, or ought to take to bear) positively on the question of whether $p$. To be open to these
questions and criticisms is to be committed to an answer to the question on which the relevant reasons bear. You can
incur such a commitment in a variety of ways: you can commit yourself to the truth of $p$ by finding convincing reasons
that you take to bear positively on whether $p$, or you might be able to cause or create such a commitment in yourself,
through some sort of action, as the Credamites seem to do, or you might simply find yourself so committed, for no
reason. But what you cannot do, what does not make sense, is just the thing you would need to do, in order to believe
at will: arrive at a commitment to a positive answer to the question of whether $p$ by finding convincing reasons that you
do not, yourself, take to settle that question, but rather take to settle some other question. By settling some other
question, you will, instead, become committed to an answer to that question, and become open to the questions and
criticisms associated with it. (I am now just repeating myself.)
Believing as an action in its own right, is not, then, a genuine possibility. Given our understanding of the commitments involved in believing and acting, nothing could be both a belief and an action. Rather, believing must stand to any attempt to believe at will as soup stands to soup-making. It is the possible product of belief-making. And so, though you may be able to bring it about that you believe, and while you may even be able to do this easily, as a basic action, you cannot believe as an action in its own right—you cannot believe at will. Moreover, this will be so for any believer.

If I have shown that believing is not an action, then I have also shown that it is not, properly speaking, the object of an intention. Thus, it turns out, the question “whether to believe” cannot be given a practical interpretation of the sort we have been attempting to give it—an interpretation on which believing is understood as an action in its own right. Rather, the question of whether to believe \( p \) should, instead, be understood to be ambiguous between the (“epistemic”) question of whether \( p \) and the (practical) question of whether to bring it about that you believe \( p \)—where the target action is belief management. Believing, itself, is not the proper object of an intention.

(In trying to create an interpretation of the question of whether to believe in parallel to the question of whether to raise your right hand or make soup, I suggested that we gloss it as the question of “whether believing \( p \) is somehow worth doing.” But, in fact, the question of whether believing \( p \) is somehow worth doing is no more practical than the question of whether making soup is somehow worth doing—by answering either question, you will arrive at a belief about an action. In answering a practical question, you would arrive at an intention.)

---

28 My conclusion is that, given our understanding of the commitments involved in action and belief, nothing could be both an action and a belief. I do not here make the stronger argument that action and belief must be understood as we understand them. I think this stronger claim likely—I think it likely that any creatures sufficiently like us will have need for concepts sufficiently like those I have here described—but I do not argue for it, here. (When discussing my argument, Setiya says, “Unless we can explain why [one’s action] cannot be subject to plural standards of justification [epistemic and practical], we cannot rule out the intentional forming of belief.” (40). But I disagree. I mean to rule out the possibility of intentional formation of belief with the weaker claim, that action is not subject to epistemic standards.)
Before turning to Setiya’s criticisms, I would like to consider why my account provides anything like a satisfying answer to the puzzle about believing at will. One might be tempted to think that the result is rather thin: all I have shown is that you cannot (intend to) believe as an action in its own right as opposed to (intend to) bring it about that you believe; any case that looks like the former (such as the case of the Credamites) must be interpreted as the latter. But I have granted that you might bring it about that you believe with ease, at least in certain limited circumstances. So it might seem that the only real barriers to believing at will were already uncovered by Bennett and Williams, who noted that, because beliefs purport to represent reality, any Credamitean capacity to bring ourselves to believe will have to be limited and circumscribed in the ways that Bennett detailed.

However, while someone who is simply interested in engineering beliefs in himself and others might find everything of interest in the limitations identified by Bennett, it seems to me that both the original intuition and much of the philosophical interest in the topic are not accounted for by noting the ways in which the Credamites are constrained.

The topic is of philosophical interest in part because of the interest of the target notion: what do we mean by “at will,” when we deny that we can believe at will? I have suggested that we mean something other than “as a basic action” or “easily, in almost any circumstance,” or “readily, in favorable circumstances,” or “whenever one sees fit,” or “according to your own judgment.” Rather, the intuition that we cannot believe at will amounts to the intuition that we cannot believe in the way we ordinarily act—in the way we raise our right hand or make soup. It amounts to the intuition that believing is not an action, not the proper object of an intention. So to say that something is “voluntary,” in this context, is to say that it is an intentional action, the proper object of an intention. (Though this a regrettably broad use of an otherwise useful word, I do think this is how “voluntary” is used, in this debate and some others.)
Moreover, we are now in a position to understand the features of intentional action that lead us to want to identify the class of the voluntary with the class of the intentional: When you intend to do or say something intentionally, you, in some sense, settle for yourself the question of whether to do or say it—a question that mentions the action in question, under the description under which it is intentional or voluntary. That is to say, when we act intentionally, we are, in some sense, the cause of our own representations, the cause of that which we represent. But, if we act intentionally by settling for ourselves the question of whether so to act, we will enjoy both a certain kind of awareness and a certain kind of discretion with respect to our intentional actions—a kind of awareness and discretion that we lack, with respect to believing and other activities that are not voluntary.

When you intend to do something or do something intentionally, we can say that you are, in a certain sense, aware of what you intend to do or what you do intentionally, at least insofar as you have (in some sense sense) settled for yourself the question of whether to do that thing: you have settled a question that represents the action, under its intentional description. Thus, there is something right in thinking that you must, in some sense, have that description in mind.29

Moreover (and, in the present context, more importantly), you enjoy a kind of discretion over that which you intend to do, or that which you do intentionally, in that you can settle the question of whether to act, like any question, for any reason you take to bear sufficiently upon it. You can decide to raise your right hand, or make some soup, or run for office, in order to win a bet, please your sister, make a point, or make a joke—you can act for any reason you take to count sufficiently in favor of so acting. In contrast, we have seen that you cannot believe for any reason you take to count sufficiently in favor of believing. You cannot believe $p$ in order to win a bet or make a joke.

---

29 This fact, it seems to me, can account for at least much of, perhaps all of, the intuition that acting intentionally involves knowing what you are doing in some special, “non-observational” way. I should stress that this is an extremely non-standard understanding of awareness. It has nothing to do with consciousness or phenomenology. Those involved in self-deceived actions or acting from unconscious motives are, in the present sense, aware of what they were doing—simply because they are doing it intentionally, on purpose. I return to this below.
You can, for those reasons, only bring it about that you believe. You can believe \( p \) only for reasons that you take to show that \( p \).

I submit that these features of awareness and (especially) discretion account for our sense that hand-raising, soup-making, and ear-wiggling (with our without the use of hands) are all voluntary, while believing is not. You can raise your right hand, make soup, or wiggle your ears for any reason that you take to count sufficiently in favor of doing so, but you cannot believe for any reason that you take to count sufficiently in favor of believing. When you take yourself to have only extrinsic reasons, the best you can do is to bring it about that you believe.

Having identified the target notion of voluntariness, we can start to see, more broadly, what can and cannot be done at will—what is and is not discretionary, or the proper object of an intention.

The class of things we can do at will, of the voluntary, can be seen as falling between two other classes.\(^3\) On one side fall those things that cannot, themselves, done for reasons at all. So, e.g., not only soup and emissions, but also forgetting and falling in love seem not to be the kind of thing that can be done for reasons.\(^3\) And, if something cannot be done for reasons at all, then it certainly cannot be for any reason that you take to show it worth doing—though, again, you could bring it about. You can, with some cleverness and some luck, make soup, generate emissions, or bring it about that you forget or fall in love.

---

\(^3\) To avoid confusion, it may be worth noting the existence of what I would call mental actions: states of mind that one adopts intentionally, by deciding to do so. For example, remembering where you parked your car, or imagining a six-sided figure, or figuring out which way is north, can, often enough, be done intentionally, for any reason that you take to show it worth doing.

\(^3\) I do not have a clear understanding of why forgetting cannot be done for reasons. Remembering, I think, can be intentional, though often it is not. And people can forget on purpose. But it seems that, though forgetfulness can be motivated, what one must intend to do is intend to make yourself forget. Perhaps this has to do with the incompatibility of forgetfulness and awareness, even of the non-psychological sort I mentioned above. Falling in love, insofar as it has anything to do with reasons, seems to me to have a structure similar to the structure Scanlon describes for valuing: it is a pattern of taking yourself to have specific reasons to do other, specific things, but the pattern itself is not something done for reasons (though, again, there are reasons for creating or eliminating the pattern, in one’s life). Cf. “Value” in T. M. Scanlon, *What We Owe to Each Other* (Cambridge, MA: Harvard University Press, 1998).
On the other side lie those things, like believing, that are done for reasons, but nonetheless cannot be done for just any reason that you take to show them worth doing. Notice that such activities will be ones for which there are, not only bad reasons (as, say, appeal to astrology provides a bad reason for believing), but also, additionally, reasons of (what have been called) the wrong kind. We can find ourselves with reasons that, on the one hand, genuinely count in favor of these activities—reasons that show them genuinely useful, attractive, convenient, etc.—but which also are not to be the right kind of reason for thing in question. These are what I have called extrinsic reasons.

I believe it is the fact that we can encounter the wrong kind of reasons for belief that generates our puzzlement about believing at will: even though believing is the sort of thing done for reasons, we can find ourselves with reasons that genuinely count in favor of believing, but which are not reasons for which we can believe. This can seem surprising and puzzling—in a way that it is neither surprising nor puzzling either that bringing yourself to believe should take some work or that such work could be made much easier with the right equipment.

Philosophers have had difficulty saying what makes reasons be of the wrong kind. I have argued elsewhere that reasons of the right kind (are taken to) support the commitments constitutive of (one could say, characteristic of) the activity in question, while reasons of the wrong kind do not support those commitments, but nonetheless (are taken to) count in favor of the activity. That is to say, reasons are of the wrong kind when they do not bear or are not taken to bear on the question proper to that activity (for believing, the question of whether \( p \)), but but rather (are taken to) count in favor of the activity in some other way (by showing it useful, convenient, or flattering, e.g.).


33 Hieronymi, "The Wrong Kind of Reason."
Notice that, in the case of voluntary, intentional action, the question proper to the activity just is the question of whether there is anything that counts in favor of doing it. Thus, in this case, we will not be able to draw a distinction between the right and wrong kind of reasons: any reason that counts in favor of the action will also support the commitment characteristic of it.

However, other attitudes or activities (such as believing) involve commitments of different sorts. They reveal our answer to a question (or questions) other than the question of whether to do them or to bring them about. So, resenting, fearing, trusting, and being proud of something or someone all reveal our take on that person or thing—they all involve commitments. And, in each case, we might reflect on our situation and find ourselves with reasons for wanting to have the attitude that do not support the required commitments—we might find ourselves with extrinsic reasons for resenting, trusting, or being proud. Thus, these attitudes or activities will be subject to a wrong kind of reasons problem. They will also be non-voluntary for the same reason that believing is non-voluntary: if we find ourselves with the wrong kind of reason, we may, with luck, be able to bring them about, but we will not be able to do them at will.

So the fact that beliefs purport to represent reality does, in the end, have a role to play in explaining the non-voluntariness of believing. But, perhaps surprisingly, the role it plays does not especially depend on belief’s concern with truth. Nor does it have to do with any required opacity or forgetfulness. Rather, the involuntariness of believing is accounted for simply by the fact that believing is committing to an answer to a question other than the question of whether believing is, itself, worth doing. It just so happens that the relevant question is whether something is so. But even

---

34 As noted above, I think certain attitudes, like loving or valuing, are not themselves done for reasons, though they do reveal our take on their object. They are not themselves done for reasons because they are rather a collection of commitments, each of which is adopted for reasons.

35 I would add to this class intentional actions under certain descriptions—under descriptions that pick out a subset of the possible reasons for so acting, such as helping kindly or attending spitefully. One can have reasons for helping kindly that are not the reasons that would qualify the action as kind. These are the wrong kind of reason for a kind helping action.
intending reveals our take on something other than the question of whether intending is, itself, worth doing. It reveals our take on the question of some action is worth doing. And intending is also not voluntary, in the relevant sense.

The non-voluntariness of intending can be hard to see, because the discretion we enjoy in acting often masks our lack of discretion in intending: reasons that show it good to intend to x can very often be sufficient reason for x-ing—you can act so as to intend. For example, the fact that you would be hurt if I did not intend to attend your party can be my reason for attending. By deciding to attend, I therein intend to attend, and thereby avoid your hurt feelings. But it will not always be the case that reasons that one takes to count in favor of intending to x are also be reasons one takes to show x-ing itself is worth doing. Perhaps you have no intention of marrying your partner, and he or she is distressed about that fact. Perhaps you generally like to avoid his or her distress, and, in fact, would be happy to house the desired intention, so long as you did not actually have to endure the ceremony and enter the legal relationship. Perhaps you even have, at your disposal, some means of ensuring that the wedding will never occur. Even so, you will not be able to form the desired intention and so avoid your partner’s distress. You will form the desired intention only if you settle the question of whether to act, and, so long as you do not think that his or her distress reason enough to marry, you will not settle that question for that reason. These uncooperative cases show that you can no more intend at will than believe at will.36

So, again, the constraint on voluntariness does not depend on belief’s concern for the truth. It is rather accounted for by the fact that believing is among the states of mind that are or represent one’s answer to a question other than the question of whether to adopt that very state.\(^{37}\)

We have discovered what I would call an illusion of reflection: our ability to reflect on our own evaluative activities leads to the illusion of a shortfall in our powers. We are evaluative creatures who answer questions. In fact, we can answer any question of which we can conceive for any reason we take to bear sufficiently upon it. We also are able to think about such activities, in ourselves and in others. We categorize them as believing, intending, resenting, trusting, etc. Further, we can notice that certain commitments or states of mind are beneficial, useful, important, inconvenient, ugly, attractive, or an imposition. Thus, reflecting upon ourselves, we can then take ourselves to have reasons for having certain states of mind which are of the wrong kind—they are not reasons that support the required commitment. In such a case, we cannot form the desired state of mind for the reasons at hand. We will have to bring it about—typically by taking steps of self-management. It is puzzling to find ourselves with reasons for something that is itself the kind of thing done for reasons, and yet unable to do that thing for the reasons we have. It can seem that we have encountered an inability. And yet there is no shortfall in our powers. The purported “inability” is simply the inevitable upshot of being able both to answer specific questions and to think about one’s answers.

---

\(^{37}\) Return, briefly, to Williams’ example of blushing. It is, I think, a very poor example. Blushing is not the same as simply have a red or flushed face (upon finishing heavy exercise, my face is red, but I am not blushing). If we set blushing aside and consider simply flushing your face, then, as with ear-wiggling, the difference between doing it “at will” and “making yourself” do it is simply the difference between doing it “just like that,” as a basic action, and doing it in some more complicated way, “by a route.” But blushing, understood (as Williams seems to understand it) as a response to some kind of emotional discomfort, could not, I think, be done at will, any more believing. You could, at most, make yourself blush, by making yourself embarrassed or uncomfortable.
Turning, at last, to Kieran Setiya’s criticisms of my earlier arguments. Setiya claims that my argument proves too much and that I have overlooked this fact because I have overlooked the distinction between believing and forming a belief. He points out that, if my arguments succeed, they would show that we cannot even intend to form a belief—but that, he thinks, we clearly can do, “as when I intend to form the belief that p next month and have yet to reflect on how” (39).

I do not think my arguments prove too much, even if we grant the distinction Setiya urges. However, I also believe we should resist the distinction between believing and forming a belief, at least as understood by Setiya. I will first attempt to say why I would resist the distinction, as understood by Setiya. However, because my remarks on this topic will be inconclusive, I will then adopt Setiya’s distinction and argue that, even granting his picture, my arguments do not prove too much. It will be apparent, throughout, that our disagreement about believing at will reflects much larger disagreements about action, intention, and belief.

Setiya understands the distinction between believing and forming a belief as an instance of a general metaphysical contrast, important to action theory, between what he calls “states, like being tall, and things can be finished or completed and in that sense done... like walking, digesting, and growing.” The contrast, he says, “corresponds to the grammatical notion of perfective aspect.”

Setiya asserts both that his contrast is exhaustive (“what can be instantiated by an object can be instantiated perfectively, like walking, digesting, and growing; or it is state, like believing, desiring, and being tall”) and that the object of an intention is never a state, but always something that can be finished or completed and in that sense done.

A quick terminological point: I have reservations about Setiya’s way of characterizing the latter class, as “things that can be be finished or completed and in that sense done,” because I think that not

---

38 Setiya, “Believing at Will,” 38.
39 Ibid.
all doings—that is, not all activities, in particular, not all activities done for reasons—are things that can be done in the sense Setiya specifies: things that can be finished or completed. In particular, I think believing is an activity done for reasons, though not something that can be finished or completed. Thus, from my point of view, Setiya’s word choice (“things that can be... done”) courts confusion. To guard against it, I will give to those things that can be finished or completed the admittedly unlovely label “potential completions.”

Notice, next, that Setiya has, here, provided his own, distinct limitation on the proper object of an intention: only potential completions are possible objects of intention. This may be so, and it may provide us additional insight into the class of the voluntary. If believing (which I take to be an activity) is not a potential completion (as I would agree), we would have another argument for why we cannot believe at will.

Setiya does not pursue this. Instead, he takes the fact that believing could not be a potential completion to show that the difficulty with believing at will must be a difficulty with forming a belief at will. So he thinks that I should be arguing, not against the possibility of believing at will, but against the possibility of forming a belief at will.

In reply, I admit to moving freely between “believing” (which sounds like an activity) and “belief” (which sounds like a state) and to neglecting entirely the locution “form the belief” (which sounds like something that could be completed). However, I do so not because I have missed or overlooked an important metaphysical contrast. Rather, I find this contrast problematic in the case of belief.

On Setiya’s picture, beliefs are states, like being tall (perhaps they are standing dispositions, like being soluble), while the object of intentions are potential completions, and this contrast is

---

40 Setiya, e.g., says that “the basic object of intention is never a state, but always something that can be done, the sort of thing of which we can ask why someone did it and evaluate his reasons.” But here I think he is trading on sense of “done” that he first indicated, something that is a potential completion, and what I think a broader sense—something for which one can be asked for one’s reasons.
exhaustive. It is thus very tempting to think that, on Setiya’s picture, any activity that could be done for reasons must be a potential completion. If this were so, then we would form beliefs, or revise beliefs, for reasons—forming and revising are potential completions—but believing would not, itself, be an activity done for reasons. It would be, rather, a state that we create in ourselves, for reasons. Having once created the belief, we would be done, finished with the part of believing done for reasons. (Like God on the seventh day, we would rest.) Thus, on this picture, we would relate to the beliefs we have formed in something like the way we relate to the (completed) documents stored on our hard drive: we once created them, and they now sit in memory. Unlike the documents on our hard drive, our beliefs are also dispositions to on-going thought and behavior. But they would affect our on-going thought and behavior, it seems, not as our own present activity, but rather as standing dispositions that we once authored (and perhaps somehow stand ready to alter).

This picture (which is not obviously Setiya’s, as I will explain below), is, I think, mistaken. I hope the ways in which it is mistaken start to show simply in its statement. The picture forces belief and believing into categories (state or completable process) into which believing does not easily fit (and into which, I think, there is no antecedent reason to think it should fit), and so distorts our relation to our beliefs and to their consequences.41 My beliefs do not sit in my mind as last week’s lecture sits on my hard-drive, recording what I once thought. My beliefs are rather my present, on-going take on what is so. When I revise a belief, I do not—or, typically I do not—notice that it no longer says what I think and then do something to change it. Rather, as I think about what is so, my beliefs therein change.42 Moreover, the current affect of my beliefs on my thought and behavior is not the indirect consequence of some past activity of mine, via a standing disposition, an affect which I

41 Matthew Boyle has been arguing this point. See Matthew Boyle, "'Making up Your Mind' and the Activity of Reason," (ms).

42 One might think I am being unfair: in insisting that believing is a state, rather than a doing, one need not imagine that the state is one that you must revise by thinking about and acting upon, like the document on your hard drive. This may be so. But then we need some account of the interface between the state and my reasoning: how is it that I affect my dispositions; by what kind of activity do I install and modify my beliefs? I press this question in Hieronymi, "Two Kinds of Agency," 156–7.
might have once predicted, when I formed the disposition, and might now notice and perhaps do something to change. My beliefs are not like the “out of office reply” feature that I set up for my email account: they are not something I create and then allow to affect things on my behalf. Rather, the affect of my beliefs on my on-going thought and behavior is an aspect of my current, on-going activity.\(^\text{43}\) Similarly, defending and giving my reasons for my presently held beliefs is not like defending or giving reasons for a past, completed activity. If you ask why I believe the butler did it, you are not looking for an account or justification of some past, completed action of mine.\(^\text{44}\) You are rather looking, now, for a case to be made for his guilt—a case to support my present stand on this question. And my answer would be no less an answer to your question if I came up with new reasons, on the spot.

I hope the above collection of thoughts gives some sense of my misgivings of any picture on which a belief is a state in contrast to an activity done for reasons.\(^\text{45}\) However, as noted above, it is not obvious that the picture is Setiya’s. This is because Setiya allows that a belief, which is a state, can be “for reasons.” He does so by allowing that one can be in the state of believing \(p\) for reasons \(x, y,\) and \(z\). What remains unclear to me is whether or not the state of believing for \(p\) for reasons \(x, y,\) and \(z\) is itself something that one does for reasons \(x, y,\) and \(z\) (even if not something one could complete), or whether it is instead somehow more passive. That is, it is not clear how Setiya will understand believing as something we do (though we do not complete) for reasons. So, at best, an

\(^{43}\) This is so, I think, regardless of whether I approve of them or endorse them or am able to dissociate myself from them, and regardless of whether they are available to consciousness. I am here working against what are sometimes identified as “real self” views of agency, views which identify the agent with some subset of her beliefs, desires or values.

\(^{44}\) In fact, it would be hard to formulate your question, if we tried to understand it this way. You might ask, “How did you come to believe that the butler is guilty?” But you would, it seems, then be asking for something like what Anscombe calls a “mental cause,” and you might miss my reasons entirely. (See G. E. M. Anscombe, *Intention* (Oxford: Blackwell Publishing Co., 1957).)

\(^{45}\) I do not take these thoughts to be conclusive. In fact, I am currently engaged in a project attempting to provide further support for the claim that beliefs are not only states but also activities.
important category (things that we do for reasons but that are not completable) has been left unarticulated.\footnote{In correspondence, Setiya has said he is willing to grant that a state like believing may bring with it certain sorts of commitments, and that perhaps, given those commitments, we should call that state an activity. I am in favor of doing so, and in the process of arguing (in different work) that we should. However, it is not clear to me how to fit these claims with the rest of Setiya’s view.}

In any case, I have not simply overlooked the metaphysical contrast to which Setiya points, but, rather, I find it at best unhelpful, in the present case. Again, the contrast forces belief and believing into categories (state or completable process) into which believing does not easily fit (and into which, I think, there is no antecedent reason to think it should fit), and so seems to me to distort our relation to our beliefs and to their effects. I hope I have given some sense of my misgivings. But, rather than pursue these misgivings further, I will adopt Setiya’s picture to reply to his criticism from within it.

Again, according to Setiya, \textit{believing} is not the kind of thing that could be the object of an intention, but \textit{forming a belief} could be. So he thinks that, to show why we cannot believe at will, I must show why we cannot form a belief at will. But he believes that my argument, applied to forming a belief, proves too much: it would show that we cannot even \textit{intend} to form a belief, which, Setiya thinks, we can do.

To assess this, we need to better understand what it is to form a belief. Since forming a belief that $p$ is the part of believing $p$ that Setiya thinks is done for reasons, I will understand it to be answering for yourself the question of whether $p$—the thing you do when you find convincing the reasons you take to bear on whether $p$.\footnote{Setiya glosses “forming the belief that $p$” as “becoming sufficiently confident that $p$” (45), where that “becoming” may or may not involve duration (47). Though I would resist the passive construction, I take “becoming sufficiently confident that $p$” to be what one does, when one answers for oneself positively the question of whether $p$.} (I am, for the sake of argument, granting that answering a question is a potential completion, though I think this a mistake.) Importantly, answering the question of whether $p$ contrasts, here, with other activities that might also be grouped
under the head “forming a belief,” such as directing your attention to the question of whether \( p \), conducting an investigation as to whether \( p \), undertaking some deliberation about whether \( p \), reflecting upon \( p \), or causing yourself to believe \( p \).\(^{48}\)

But once we distinguish forming a belief from these other, nearby activities, I am happy to say that we cannot intend to form a belief (though we can intend to deliberate about whether \( p \), or investigate whether \( p \), or reflect upon whether \( p \)): we cannot intend to answer the question of whether \( p \), where answering that question is now understood as an ordinary action in its own right. The earlier arguments apply.\(^{49}\) So I deny my argument proves too much.

Because Setiya believes my arguments prove too much, he considers what might have gone wrong. He offers two possible diagnoses, each of which suggests a more expansive notion of intention than I would adopt. I will not examine the diagnoses in detail, simply because I believe that the larger issues remain mostly off-stage. I take an intention to \( x \) to be the state of mind one is in, when one is committed to a positive answer to question of whether to \( x \). As noted in the previous section, this at once accounts for our sense that intentional actions (i.e., the proper objects of intentions) are voluntary and rules out intending to believe (or, intending to form a belief). In contrast, Setiya characterizes intention as “the kind of state that motivates one to do what can be done, and guides it to completion.”\(^{50}\) This characterization does readily connect what is intentional and what is, in the sense highlighted in the previous section, voluntary. (So, Setiya sees no difficulty in intending to

---

\(^{48}\) Matthew Boyle has also insisted on these contrasts. See Boyle, "Making up Your Mind' and the Activity of Reason."

\(^{49}\) You can answer the question of whether \( p \) for reasons you take to bear on it or, perhaps, for no reason, but, in neither case would have intended to answer the question. You would form an intention to answer the question only if you take yourself to have only extrinsic reasons for answering it. But you cannot, by finding these convincing, answer the question. Rather, you will, by finding the extrinsic reasons convincing, answer the question on which you take the extrinsic reasons to bear: namely, the practical question of whether answering the question of whether \( p \) is in some way good to do. Thus you will, by finding your reasons convincing, believe that answering the question of whether \( p \) is good to do. You might also form an intention to bring it about that you answer the question of whether \( p \), through, say, directing your attention in certain ways or conducting an investigation or taking time to deliberate. But you will not, therein, answer the question of whether \( p \).

\(^{50}\) Setiya, "Believing at Will," 39.
form a belief, e.g.). (I suspect that the capaciousness of his notion of intention is what allows Setiya
to do largely overlook those activities done for reasons that are not, on either of our accounts, the
proper object of an intention.) The larger issue is which picture of intention is ultimately most
satisfying, intuitive, plausible, or useful.

**SETIYA’S ARGUMENT**

Turning, now, to Setiya’s own arguments. Setiya argues that believing at will is not possible on the
basis of two premises together with a claim about what it would be to believe at will. Because Setiya
thinks we can intend to form a belief, he needs a more restrictive account of what it would be to
believe at will. According to Setiya, to believe at will would be to form a belief intentionally
“irrespective of its truth,” that is, “to form the belief that \( p \) by intentional action, believing
throughout that, if one were to form that belief or to become more confident that \( p \) intentionally,
one’s degree of confidence or belief would not be epistemically justified.”

The first premise to the conclusion that we cannot, in this sense, believe at will is about belief:
“It is impossible to believe that \( p \) or to be confident that \( p \) while believing that this degree of
confidence or belief is not epistemically justified.” The second premise concerns self-knowledge
and intentional action: “If \( A \) is doing \( x \) intentionally, he believes he is doing so, or else he is doing \( x \)
by doing some other intentional action, in which he does believe.” From the second premise, Setiya
concludes that, if one were to intentionally form a belief “irrespective of its truth” one would have
to believe that one were doing so. But then one would violate the first premise: one would believe
while believing one’s belief unjustified. The result is that “believing at will without a failure of
attention or logical confusion would require a lapse of self-knowledge, an ignorance of what one is
doing intentionally that conflicts with it being intentional. This is the sense in which, and the extent
to which, it is impossible to believe at will.”

---

51 Ibid.: 42.
52 Ibid.: 47.
confusion is to capture those cases in which one fails to infer, from the fact that one is intentionally believing $p$ without evidence, that one’s confidence is not epistemically justified.)

Although this is an extremely elegant argument, I am disinclined to believe either of its premises. They both seem to me to make the tempting mistake of confusing what is ideal or rational in a mind with what is possible for it—they overlook how messy and irrational actual minds can be.

The first premise encounters the clearest difficulty. I see no reason to think it impossible for someone to believe something while also believing her belief unjustified (and, to be fair, Setiya admits that he is at a bit of a loss in supporting this principle). In fact, I don’t see why it should be impossible to do this “in full consciousness” (though, insofar as it conflicts with a standard of rationality, it will not be possible to do “in full rationality”). Perhaps, stuck in an affair, you find yourself convinced that your lover will leave his or her spouse, while fully aware your belief is ungrounded and unjustified. Or, perhaps you believe that God exists and is good, while also believing both that your faith ought to be based on conclusive evidence and that the evidence is inconclusive. To quote Harry Frankfurt on a related topic, “That may not be a sensible way to go about things. Nevertheless, it can be done.”

Setiya supports this first premise by suggesting that, in the kind of case imagined, we should decline to call your state of mind a belief. Rather, “apparent violations of this principle are better understood as cases in which one has a nagging thought or tendency to act as if $p$, even though one does not believe it.” I find this interesting but unconvincing. Notice that, to think your (purported) belief is ungrounded or unjustified, you need not think it is false. Nor need you have any tendency to doubt it. You need only to think that there is no justification for it, nothing showing it to be so (and also, perhaps, that there ought to be some justification, that this is not a

53 Harry Frankfurt, "Disengaging Reason," in *Reason and Value: Themes from the Moral Philosophy of Joseph Raz*, ed. R. Jay Wallace, et al. (Oxford: Oxford University Press, 2004). To clarify: these things can’t be done at will, but they are possible states to find oneself in, and are states of one’s own doing.

54 Setiya, "Believing at Will," 43.
cases in which believing without ground is appropriate). But I do not see why this should require us to withhold the attribution of belief, given that you think \( p \) is true, and think and act accordingly, with no tendency to doubt it.

(In contrast, I would make just the claim Setiya makes, were someone to say that some purported believer of \( p \) is not committed to the truth of \( p \)—not vulnerable to questions and criticisms that would be satisfied by reasons that bear on the question of whether \( p \). To the degree that we deny this vulnerability or commitment, to that degree, I think, we should decline to attribute belief and instead attribute some other sort of state of mind (a persistent thought, a nagging illusion, a vivid representation, etc). So, while Setiya thinks believing requires a kind of internal coherence or lack of self-condemnation on the part of the believer, I think that it requires, instead, a certain kind of answerability for one’s state of mind.)

The second premise also seems problematic, though the issues here are more difficult. The second premise seems belied by the possibility of self-deceived, unconscious, or subconscious actions. You might, after all, be showing off your extensive knowledge, laying a guilt trip on your children, getting revenge, or flirting with your subordinates, and doing so intentionally, without being consciously aware that you are doing so and without there being any reason for us to attribute to you a belief that you are doing so—beyond, perhaps, the mere fact that your action is intentional. One might simply deny such actions the title “intentional” (as one might deny them the title “autonomous action” or “full-blooded human action”\(^{55}\)), but this seems to me an unhelpful verbal stipulation. Such actions are surely done on purpose, for a reason, with a plan, and, I think, open one to the same kinds of questions and criticisms as (other) intentional actions. It seems to me much better to

---

say that they done intentionally, and then perhaps to add that intentional action should satisfy certain ideals that these fail to satisfy.\textsuperscript{56}

Setiya supports his second premise by appeal to Anscombe’s observation that, if you do not know that you are $x$-ing, the question “Why are you $x$-ing?” is refused application, and the applicability of that question marks intentional action.\textsuperscript{57} Thus, it seems, if you do not believe you are $x$-ing, you are not $x$-ing intentionally. I find this appeal to Anscombe interesting, as her work, including, centrally, this observation, has inspired my own approach to action.

In the kinds of cases Anscombe imagines, a person shows herself not to have been $x$-ing intentionally by revealing that she did not know she was $x$-ing. (“Why are you standing on the hose?” one person asks. “Oh, I did not realize I was doing so!” the other replies, thereby showing that she was not intentionally standing on the hose.) Anscombe’s cases are compelling. Yet I find it perverse to deny the title “intentional” to the sorts of self-deceived or unconscious activities mentioned above (again, unless one simply uses “intentional” as a term of art, stipulating that one reserves it for cases that do not involve self-deceit or unconscious motivations).

I believe both intuitions can be preserved. Anscombe’s observation is meant to highlight an interesting and puzzling feature of intentionally $x$-ing: when one intentionally $x$’s, it seems that one knows what one is up to “without observation.” I have advocated a particular account of this interesting and puzzling feature: whenever one $x$’s intentionally, or intends to $x$, one has, in some sense, settled the question of whether to $x$—that is, one has settled a question that represents one’s

\textsuperscript{56} Setiya allows one to act intentionally without believing that you are doing so in cases like Davidson’s carbon copier. (In Donald Davidson, ”Intending,” in Essays on Actions and Events (Oxford: Oxford University Press, 1980).) Setiya formulates this by saying that one must “be doing some other intentional action, in which [one] does believe.” But I take it that the idea is better stated on page 47, when he says “one must believe that one is doing it as an intentional action or that one is taking further means”—that is, I take it that the requirement is not merely that one must be doing something that one believes one is doing (pressing hard on the first copy), but also that one must believe that doing so is a means to the thing that one does not know one whether one is doing (making ten carbon copies). So, the mere fact that one believes that one is making a funny joke would not be sufficient to secure for Setiya the claim that one is intentionally flirting with one’s subordinate, unless one also believed that making the joke was a means of flirting with one’s subordinate.

\textsuperscript{57} See Anscombe, Intention.
action under a certain description. This fact both accounts for one’s vulnerability to the question “Why are you x-ing?” (this why-question looks for the reasons for which one would have settled the question of whether to x) and accounts for our sense that you should, in some way, know what you are up “without observation:” if you have settled a question that represents your action under a certain description, you should, in some sense, have that description in mind. Thus, if a given description comes as a surprise to you (“I did not realize I was standing on the hose”), that will typically show that you did not settle the question of whether to do that thing under that description, and so will typically also show that your action is not intentional under that description.

However, one need not insist that every case is typical, and so need not insist that, unless someone is consciously aware that she is x-ing, or unless there are independent grounds for attributing to her a belief that she is x-ing, she cannot be x-ing intentionally. Rather, we can allow the possibility that someone is doing something intentionally—that she has, in some sense settled the question of whether to do that thing—even when she is otherwise unaware of doing it or sincerely denies knowledge of her action. Perhaps, to the extent that the action is intentional, we will, on that basis, impute to her a belief that she is so acting, or at least some sort of unconscious awareness of what she is up to—however, the imputed belief or awareness will be just as self-deceived or unavailable as the intentional action. So, I doubt that (what is sometimes called) practical knowledge can place any independent constraint on what is intentional.

Some compare and contrast may be useful.

While I claim that believing cannot be an action in its own right, because the commitment required for belief could only be the product of a distinct, managerial action of bringing yourself to believe, Setiya claims that bringing yourself to a belief irrespective of its truth could not be intentional, because intentional actions require a belief that would be, in such a case, incompatible with the target belief.
On my view, we are creatures with a take on the world, both on what is so and what we shall do. Our take on what we shall do constitutes our intentions; our take on what is so constitutes our beliefs. We are social, non-solipsistic creatures, sharing a world of limited resources, and we expect one another to make sense—that is, we both rely on the assumption that others will make sense, in our predictions, and we demand that others do make sense, in our interactions with them. Beliefs or intentions that are unreasonable or incoherent, in either their object or their grounds, are subject to criticism. We are also reflective creatures, capable of thinking about ourselves and our own attitudes. We can notice that certain attitudes, certain aspects of our take on the world, may be more or less attractive, convenient, useful, irritating, or debilitating. We can thus notice that we have reasons that count in favor of or against housing these attitudes, which are nonetheless of the wrong kind of reason for them. By finding these extrinsic reasons convincing, we will not change the attitude they so to be good or bad. Rather, by finding them convincing, settle the question on which we take them to bear—the question of whether a given attitude is good or bad to have. However, once we have decided that it is good or bad to have this or that attitude, we can, or perhaps with the right technological advances we could, bring it about that we arrive at the desired state of mind. With the right equipment, we might do so easily. The only thing we cannot do is to settle a question by finding convincing reasons that we do not take to settle it. Nothing could count as doing that. But, that is just what we would have to do, in order to believe at will—in order for believing to be an action in its own right.

On Setiya’s view, the mind encounters stronger restrictions. It is not only irrational or unwise or untypical, but impossible to believe while believing your belief unjustified or to act intentionally without believing you are doing so. Bringing yourself to a belief, irrespective of its truth, is impossible, because it runs afoul of these stronger restrictions. I suspect Setiya is drawn to these stronger restrictions because, like many, he would like first to describe the workings of the mind and then use that description to ground or explain various ideals or norms or standards or commitments.
So he would like to account for those norms or standards by appeal to the existence conditions of the states of mind that they govern. It is a common enough strategy. But, not only do I find the approach unmotivated—unnecessarily restrictive in the resources it allows for itself—but it also seems to me that those who try this route often end up saying implausibly restrictive things about when the states of mind are possible.

CONCLUSION

I have, I hope, re-presented my argument that believing at will is impossible in a way that is intuitive and brings out what I take to be important about the puzzle about believing at will. In doing so, I hope to have displayed some of the connections between my answer to this question and broader questions about the voluntariness or discretion of ordinary action, as compared with other activities done for reasons. Finally, I have briefly considered both Setiya’s arguments against my position and his own argument to a similar conclusion, and I hope I have drawn out some of the broader and more important differences between our approaches.58

58 This essay has been improved by written comments from and/or conversation with Matthew Boyle, David Hunter, Mark C. Johnson, Lucy O’Brien, Kieran Setiya, Nishi Shah, and Eric Schwitzgebel. Thanks are also due to audiences at Ryerson University and CSU Northridge. Special thanks are due to David Hunter, for his work in organizing the conference for which this essay was written, and to Kieran Setiya, for prompting the exchange.
BIBLIOGRAPHY


Boyle, Matthew. "'Making up Your Mind' and the Activity of Reason." (ms).


