Control, Beliefs, and Moral Responsibility
Corinne McCarthy, Class of 2009
Philosophy Honors Thesis
Submitted May 6, 2009
Abstract

Typically, we assume that an agent can be responsible for some thing only if he can do that thing just by willing to do so. Thinking this way leads to an interesting question regarding our responsibility for our beliefs. That is, we tend to hold individuals responsible for their beliefs and yet, they are things that cannot be acquired at will. In an effort to answer this question, I consider views of philosophers such as Angela Smith who have recently argued that we do not need to be able to control things such as our beliefs in an “at will” sort of way in order to be responsible for them. I point out that we have good reason to doubt that we must be able to do things at will to be responsible for them since we often lack this sort of control over what we consider the paradigm for which we can be held responsible. Expanding on Smith’s view, I argue that we can be held responsible for our beliefs because we exercise a robust form of agency in coming to believe $p$.

Introduction

Very often, we tend to make judgments about our own and other individuals’ beliefs. We may for example judge another person’s beliefs to be unreasonable, based on insufficient evidence, or simply foolish. Sometimes we are even prone to respond to individuals’ beliefs in a stronger sense. More specifically, we commonly hold individuals morally responsible for their beliefs. That is, we often find individuals to be open to moral credit or discredit on the account of their beliefs.\footnote{In order to avoid any confusion, it is important to briefly outline the notion of responsibility that I take to be at question. To elaborate, we tend to use the expression ‘A is responsible for X’ in different ways (Oakley 124). For example, we sometimes say that A was responsible for X in the sense that A did nothing more than that A caused X. This sense of responsibility however need not carry any implication of credit or blame and is often associated with actions and events rather than agents as causes (Wolf 40). Thus, we might say, ‘the bent axle is responsible for the noise that the car is making,’ which means no} Moreover, we tend to hold each other responsible for our beliefs in a direct sort of
way. To elaborate, we do not regularly hold each other responsible for our beliefs merely as features of ourselves like our health that we can manage or affect through our actions (Hieronymi 2008: 357). Rather, we tend to hold each other responsible for our beliefs in something similar to the way that we hold each other responsible for our actions themselves (ibid).  

Perhaps an example will more clearly demonstrate the ways in which we are typically inclined to hold each other morally responsible for our beliefs. Consider the following example. An intelligent male student believes that his attractive female colleague is intellectually inferior simply because she is an attractive woman. I take it that most of us would probably feel at least mild resentment towards the male student for holding his belief under these circumstances. This negative attitude that we would adopt seems to be different in form than the dissatisfaction that we might feel about certain aesthetically displeasing features of an individual. In fact, this attitude seems much closer in kind to the responses of hurt feelings, anger, and so on that we typically adopt in response to deliberate actions that we consider morally blameworthy. We might for example say to the male student things like, “How could you believe such a thing?” or “You ought not to have believed such a thing about your peer,” where these criticisms suggest that we take the male student to be responsible and open to moral assessment for his belief.

Footnote 1 continued: more than that the bent axle caused the noise that can be heard coming from the car. Additionally, we might use this expression to say that an agent is responsible for some thing in the sense that she is blameworthy for it. We might also use the expression ‘A is responsible for X’ to indicate that A is open to praise or blame on the account of X (Berofsky 45, Scanlon 248, and Smith 2005: 238). However, whether A is actually blameworthy or praiseworthy for X depends on its rightness or wrongness (ibid). It is views about the type of control that is necessary for this latter sense of responsibility that I want to examine in this paper.  

Some philosophers think that we can be responsible for our beliefs insofar as they are the result of activities that can be performed voluntarily. While I agree that there are certain activities that we can voluntarily perform that influence our beliefs, I think that these philosophers have missed the point of the problem that I take to be at issue. The problem arises because we take ourselves to be more directly responsible for our beliefs than say the state of our rooms that we can control by carrying out certain voluntary activities such as painting the walls, vacuuming the carpet, and so on (Hieronymi 2008: 358).
Furthermore, I presume that most of us would feel this sort of resentment even if this individual’s belief did not lead him to perform any sort of harmful action.

Although we regularly react this way to each other’s beliefs, there seems to be something problematic about holding individuals morally responsible for their beliefs themselves. The problem is predicated on the assumption that we can only be held responsible for what is under our control. This assumption seems to call into question our responsibility for our beliefs since our beliefs do not seem to be controllable in an ordinary sense. The problem is that we cannot control our beliefs in the way that we can control raising our hands, looking to the left, shaking our heads, or our other movements. If an able bodied person for example chooses to raise her arm, she can just like that raise her arm. A believer however cannot just choose to believe that it is Easter Sunday, that it is snowing outside, or that the sky is red and just like that believe these things.

A believer can also find herself with beliefs that she does not want to have, and with little that she can do about these beliefs (Feldman 1988: 238). After learning of certain facts, a believer might wish for example that she did not believe that her friend acted dishonestly. While she might be able to put the matter out of her mind, she cannot just change her belief about this topic just because she wants to or just because she thinks that she would be better off without the belief. Moreover, our inability to change our beliefs at will seems to be more than a mere

---

3 Admittedly, “choosing to believe” or “deciding to believe” may have two very different meanings. We sometimes use the expression ‘she decided to believe that p’ to mean that she formed a belief by making up her mind about what is true (Hieronymi 2006: 46). We might for example say that an individual decided to believe based on the available evidence that she can make it to the flight on time, whether an item is priced too highly, or whether it is likely to snow (ibid). I am however not referring to this sense of “choosing to believe” or “deciding to believe.” In saying, ‘an agent could not just choose to believe some thing.’ I mean that an agent could not just will to believe some thing and then immediately believe that thing without doing anything else.
psychological fact. In fact, our inability to change our beliefs at will seems to be traceable to the very nature of beliefs, as we will see later on in the paper.

In this paper, I will explain that we can indeed hold individuals morally responsible for their beliefs despite these considerations. My strategy will be to show that any such reservations about doing so arise from a common misconception about what sort of control is required for individuals to be held morally responsible for their beliefs. This sort of control does not consist in the ability to choose to believe or not to believe some thing at will. Rather, this sort of control refers to the agency that we are able to exercise in evaluating and responding to the evidence or the reasons for believing things.

I.

Commonly, we tend to think that we can only be held morally responsible for things that are sufficiently under our control. The ability to do or have some thing at will has often been put forward as a part or the whole of such control (Stocker 447). So, for example, I can be responsible for sitting, standing, or holding my breath in part because I can sit, stand, or hold my breath at will. Conversely, I cannot for example be held responsible for sneezing because I cannot sneeze at will. To clarify, I can do or have some thing at will if all that is required for it to be the case is that I will it.

Of course this concept of “at will” is not very well defined. As Pamela Hieronymi and many others have pointed out, many of the acts that we typically take to be under our control cannot simply be done at will (Hieronymi 2008: 364, Midgley 152-154, and Solomon 409-410). We cannot do these things just by choosing to do so because they depend on external factors

---

This assumption that we must do some thing voluntarily for us to be responsible for it has been traced back to Aristotle. In the beginning of Book III of the Nicomachean Ethics, Aristotle suggests that attitudes and acts “receive praise or blame if they are voluntary, but pardon, sometimes even pity, if they are involuntary” (Aristotle 53).
Consider moving a chair in a room. We need to be strong enough to push and carry the chair in order to move the chair in the room. So perhaps the suggestion is that we can be held responsible for some thing only if we can do or have it at will given a suitable background (Hieronymi 2008: 364). For example, I can be held responsible for moving a chair in a room if, given that I am strong enough to move the chair in the room, I can perform the act of moving the chair just by choosing to do so.

Additionally, others have argued that what it means for some thing to be under our control is not just that we be able to do or have some thing at will but also that we able to not do or have some thing at will (Oakley 127). Put another way, the view here is that if we are unable to refrain from doing or having some thing just by choosing to do so, then we cannot be responsible for it (ibid). Presumably, the idea is that if we are unable to refrain from doing or having some thing at will, then we are not active and responsible in regard to doing or having it. So the argument goes that I for example can only be responsible for raising my arms if I can keep them at my sides.

A proviso needs to be added here. As Justin Oakley points out, the problem is that “we can almost always avoid doing or having some thing at will in so far as we can for example knock ourselves out or commit suicide” (Oakley 127). However, the specific sense of avoidability that is conveyed by this last view cannot of course be taken to include what can only be avoided by such radical measures (ibid). This last view cannot be taken to include such radical measures because the avoidability condition would become insignificant since little if anything would be ruled out from the realm of responsibility (ibid). To summarize, the view must be that we are responsible for doing or having some thing only insofar as we could have used reasonable means to refrain from doing or having some thing.
Still, this view needs to be clarified. To elaborate, the ability to refrain from doing or having some thing just by deciding to do so does not normally seem to be an ability that we think that individuals must necessarily possess to be held responsible (Stocker 414). An individual for example cannot stop his plane from crashing to the ground when it is plummeting to the earth on the account of a failed engine. Nevertheless, he can still be held morally responsible for crashing his plane and for injuring his passengers if he for example failed to check the condition of his plane’s engine before departing. Furthermore, Michael Stocker tells us that it is often impossible once we have acted for us to prevent the foreseeable consequences of many of our actions from taking place just by choosing to do so (Stocker 415). Once an individual has for example dropped a friend’s clay pot in order to break it she cannot usually stop it from breaking unless she has very sharp reflexes. Yet, it still seems appropriate to hold such agents responsible for many of the consequences of their actions even if they cannot immediately prevent them.

Perhaps we can avoid these sorts of problems by considering how a person came to be in a position in which her doing or having some thing became unavoidable. In the case of the broken clay pot, the individual could have avoided breaking the pot had she not dropped it with the intention of breaking it in the first place. It seems appropriate therefore to hold the individual responsible for breaking the clay pot because the reason that she came to be in a position in which she was unable to keep the pot from breaking goes back to the fact that she originally intended to break the pot. If however an agent could not have avoided doing or having some thing just by choosing to do so at any time, we tend to think that it is not appropriate to hold her responsible for doing or having some thing. In fact, many individuals argue that the ability to do otherwise is a necessary condition of moral responsibility (see for example Chisholm 255-256, Fitzgerald 143, Glover 10-11, Hart 32, and Plamenatz 173).
Prima facie, defining the sort of control that agents must exercise in order to be held morally responsible for some thing in the previously discussed ways seems natural and convincing. If one day for example you leave all your obligations behind and completely out of the blue, move to another country just because you have decided to do so, you will probably be held morally responsible for just leaving all your obligations behind. If however your sudden disappearance is due to the fact that you were kidnapped and forced to leave the country, you hardly can be blamed for leaving everything behind. Most likely, you will not be blamed for leaving because you did not voluntarily choose to leave everything behind. In fact, you lacked the ability to do otherwise than leave because of the fact that you were taken away from everything against your will by your kidnapper.

To elaborate, the problem is that we seem to suppose that an agent must have freely done the thing in question for her to be held morally responsible for it. Part of what this means is that we can trace an agent’s doing so back to her own will rather than some aspect of her circumstances such as the fact that she is being constrained or hindered when she acts. So identifying the requisite control with our voluntarily choosing to do some thing is appealing because it allows us to acknowledge some of the ways in which our circumstances affect the ways that we act.5

Putting this all together, we tend to think that the ability to do or have some thing at will and the ability to not to do or have some thing at will constitute the sort of control that agents must exercise in order to qualify as morally responsible. Of course this needs a bit of clarification since by the ability to do or have some thing at will we mean the ability to do or

5 This sort of account of the control necessary for holding individuals morally responsible for things might not seem appealing to everyone. In particular, incompatibilists might be dissatisfied with this sort of account insofar as what we voluntarily choose to do can be determined by our genetic inheritance and early upbringing. Nevertheless, my point is that we commonly tend to think that this sort of control is necessary for moral responsibility given our ordinary ascriptions of responsibility.
have some thing at will given a suitable background. Moreover, we mean the ability to reasonable do otherwise by the ability to not do or have some thing at will.

II.

Applying this account of control to the case of beliefs, we can only be held morally responsible for our beliefs if we have the ability to believe \( p \) at will and the ability to refrain from believing that \( p \) at will. Unfortunately, we do not seem to have this type of control over our beliefs.\(^6\) William Alston has most probably fiercely argued that we cannot acquire propositional attitudes at will.\(^7\) His central argument takes the following simple form:

I shall merely contend that we are not so constituted to take up propositional attitudes at will. My argument for this, if it can be called that, simply consists in asking you to consider whether you have any such powers (129).

Alston goes on to invite readers to try to change a belief that they hold through an act of will and argues if they fail to do so, they should conclude that such changes are impossible. To elaborate, Alston asks, “Can you, at this moment, start to believe that the United States is a colony of Great Britain just by deciding to do so?” (Alston 263). Moreover, Alston states that

---

\(^6\) Some philosophers dating back to Descartes have suggested that we can believe at will. Descartes tells us in the Fourth Meditation that we come to have erroneous beliefs when “…our will extends further than our intellect, we do not contain the will within the same boundaries; rather, we also extend it to things we do not understand. Because the will is indifferent in regard to such matters, it easily turns away from the true and the good; and in this way I am deceived and sin” (Descartes 84). Yet, most current philosophers acknowledge that it is very hard to see how it is possible to believe some proposition at will. This view that individuals do not enjoy voluntary control over their beliefs has often been called doxastic involuntarism.

\(^7\) Other philosophers have suggested that believing at will is impossible. In his “Deciding to Believe,” Bernard Williams for example suggests that believing at will is conceptually impossible because our beliefs are things that we take to be true and that we take to represent reality (Williams 148). In other words, Williams suggests that we cannot just choose to believe any proposition that we fancy because our beliefs by their very nature purport to represent reality (ibid).
you could not even believe that the United States is a colony of Great Britain if you were sufficiently motivated or interested in doing so without deceiving yourself in a way inconsistent with the evidence (ibid). He states: “Suppose that someone offered you $500,000,000 to believe it…It seems clear to me that we have no such power” (Alston 263).

Alston goes on to explain that we also lack the ability to refrain from believing $p$ at will. When I see a tree outside of my window, for example, I come to believe that there is a tree outside of my window. Given my visual perception of a tree outside of my window, I cannot just choose to refrain from believing that there is actually a tree outside of my window at will without deceiving myself in a way inconsistent with the evidence available to me. What’s more is that this seems true of many of our other beliefs. Upon learning of certain facts, we could not refrain from believing for example that our friends betrayed us or that the population of China is larger than the population of the United States. We might able to put these matters out of our minds but the point is that we could not just refrain from believing these things.

What Alston is getting at is that we do not have the complete freedom to choose our beliefs (Hieronymi 2006: 46). To elaborate, a believer cannot believe whatever she fancies just by deciding to do so in the way that an able-bodied individual can raise her limbs just by choosing do so. Of course, this is assuming that features of the world beyond our own body will cooperate. For example, we might not be able to raise our arms just by choosing to do so.

---

8 To elaborate, believing contrary to evidence is not a psychological or conceptual impossibility. There is of course the possibility of self-deception or of wishful thinking (Naylor 430). However, it is not clear whether the beliefs produced by such means are produced at will. The problem is that we are able to induce in ourselves these beliefs by means of self-hypnosis, that is, we must engage in a process to convince ourselves contrary to the evidence that some proposition is true or not true (Naylor 430). In other words, we need to talk ourselves into believing some proposition. Moreover, self-deceived individuals and wishful thinkers are rarely described as free in their beliefs (Engel 4). Rather, we tend to think that these individuals are compelled to believe certain things by the allure of pleasure or some other desire (ibid).
because someone might be holding our arms down. Still though, it does not seem correct to say that a well functioning believer can believe that \( p \) or that not \( p \) at will given a suitable background of evidence.

Perhaps this objection should be fleshed out more clearly. To elaborate, individuals might argue that one can decide to create in oneself a belief that \( p \) and then voluntarily perform certain acts such as making a phone call, conducting an interview, taking a pill, or turning off the lights that will result in one believing that \( p \) (Feldman 2001: 81-82 and Kelly 169-170). Richard Feldman for example points out that we can decide to believe things about particular states of the world because we exercise control over those states of the world (Feldman 2001: 82). He explains that we can for example bring ourselves to believe that the lights are on or off in a room because we usually can voluntarily flip the light switch in the room that controls the lighting in the room (ibid).

With respect to a particular issue, Feldman also tells us that we might decide to create in ourselves a belief that \( p \) and then voluntarily set out on a course of action that will expose us to the relevant literature and evidence in support of \( p \) (Feldman 2001: 81). If we for example decided to create in ourselves the belief that the world is flat, Feldman explains that we might attend meetings of the Flat Earth Society, read conspiracy literature asserting that satellite photos are fake, and so on (ibid). By doing these sorts of things, we might come to believe that the world is flat.

These sort of arguments however do not show that we can believe that \( p \) or that not \( p \) at will. Even though we can voluntarily carry out these sorts of activities, we still seem to be at the mercy of our evidence. Let us reconsider Feldman’s light switch example. Although it is often within our power whether the lights are on or off in a room, our beliefs about the lighting in the
room still seem to be constrained by whether the lights are on or off in the room. The problem is that we cannot for example just choose to believe that the lights are on in a room if the lights in the room are actually off. Nor can we just choose to believe that the lights are off in a room if the lights in the room are actually on.\(^9\)

Similarly, the fact that it is within our power to expose ourselves or not to the relevant evidence does not show that we can voluntarily choose what attitude we take toward some proposition, \(p\) (Alston 279). At most, these sorts of activities show us that we can voluntarily choose to carry out certain activities in order to expose ourselves to the relevant sort of evidence or to change the state of the world in favor of some candidate or candidates for belief. In other words, the problem remains that we cannot but adopt or reject a given belief in light of the evidence we have for or against it.

More generally, we cannot believe that \(p\) at will or believe that not \(p\) at will precisely because of the type of states that our beliefs are. To elaborate, believing that \(p\) is a matter of settling the question whether \(p\) (Hieronymi 2006: 50).\(^10\) So we cannot believe that \(p\) for any set of considerations that we take to count sufficiently in favor of believing \(p\) (Hieronymi 2008: 367). Suppose, for example, I offered you $10,000 to believe that the sky is red. You may want to believe that the sky is red because it would be good for you to do so as you would win money. In other words, you may take the reward as a reason that counts sufficiently in favoring of

---

\(^9\) Also our ability to exercise this sort of control does not seem to solve our problem because many of our beliefs for which we are held responsible for involve states of affairs that are not directly under our control in the way that the lights in a room are.

\(^10\) As Hieronymi explains, “settling the question” here does not amount to consciously entertaining the question. Settling the question whether \(p\) is nothing other than believing that \(p\) (Hieronymi 2008: 361). This expression “settling the question whether \(p\)” is just to clarify what it means to believe that \(p\) (ibid). To elaborate, it seems that in coming to believe that \(p\) we have settled for ourselves the question of whether \(p\) insofar as the question whether \(p\) asks whether the world is as \(p\) would have it and we think that our beliefs by their very nature represent reality (ibid).
believing that the sky is red. You cannot however believe that the sky is red just because you think it would be good to do so.

We cannot believe the sky is red just because we think it would be good to do so because believing the sky is red is a matter of settling the question of whether the sky is actually red (Hieronymi 2006: 59-60). And reasons for which we take it to be good to believe that the sky is red do not answer the question of whether the sky is actually red (ibid). In fact, we seem to settle a different question in finding reasons that we take to show that believing the sky is red is good to do (Hieronymi 2008: 367). More specifically, we settle the question of whether believing the sky is red is good to do (ibid). And in doing so we form a second order belief about the belief that the sky is red namely, the belief that believing the sky is red is good to do (ibid).

Believing that \( p \) is a matter of settling the question whether \( p \) because our beliefs “aim at truth” (Williams 148). Unlike desire and other motivational states, believing that \( p \) entails regarding \( p \) as true or more specifically a truth already in existence (Velleman 249). The direction of fit is different because in believing, at least in the standard sense, we seek to alter our beliefs so as to accord with the ways that things actually are (Velleman 250). However, in desiring, wishing, and so on we seek to arrange the world so as to fit our desires, wishes, and so on (ibid). Since our beliefs purport to represent reality, we could not believe some thing just by choosing to do so at will. We could not for example believe that there were tomatoes on the counter just by deciding to do so at will. Rather, we would have to come upon evidence that we would take to sufficiently show that there were actually tomatoes on the counter.

To recap, assuming that a sort of “at will” control is the sort of control that agents must exercise in order to be held morally responsible for their beliefs leads us to question our responsibility for beliefs. Again, it leads us to doubt that we can be held responsible for our
beliefs because we cannot just choose to believe that \( p \) or that not \( p \). The problem is not that under different epistemic conditions, we could not believe otherwise than we actually do. Rather, the problem is that we cannot but adopt or reject a given belief in light of certain information or evidence that we encounter for or against it.

III.

Despite these sorts of difficulties, a number of philosophers have recently proposed that we can appropriately hold each other morally responsible for our beliefs.\(^{11}\) According to these philosophers, who have been identified as the attributionists, what really matters in determining a person’s responsibility for some thing is whether that thing is appropriately attributable to her (Smith 2008: 367-368). Unlike our height, our skin color, or our compulsions, the attributionists explain that we can be appropriately held responsible for things such as our beliefs that are expressive of our judgments, values, or normative commitments because we can justify them by appealing to our normative commitments, values, or judgments. So these philosophers make judgment sensitivity, character expressiveness, or judgment dependence rather than the ability to do or not to do at will the basic criterion of moral responsibility (Smith 2008: 368).\(^{12}\)

Control is not completely irrelevant on attributionist accounts. To elaborate, control is relevant insofar as it determines what can be attributed to the agent. In determining a person’s responsibility for a bodily movement, for example, it may make sense to ask whether an agent has voluntarily chosen the thing in question because this will determine whether the particular movement can reasonably be taken to express her judgments, values, or normative commitments.

\(^{11}\) In fact, these philosophers suggest that we can be held responsible for a variety of things such as omissions, emotions, and desires that are not under our control in the “at will” sense of control (Adams 21, Scanlon 20, and Smith 2005: 242-250).

\(^{12}\) Attributionist views have been identified as modern versions of what Susan Wolf has called “real self views” because they attempt to account for an individual’s responsibility or lack of responsibility for her actions and attitudes in the terms of fact that they do or do not express who she is as a moral agent (Smith 2008: 23-45).
(Smith 2008: 368). Suppose that an individual accidentally loses her footing and bumps into a table knocking over and breaking your favorite vase that was formerly sitting on the table. While the individual is causally responsible for breaking the vase, the attributionists claim that the individual should not be held morally responsible for breaking the vase because her action was not reflective of her judgments, values, or normative commitments.

Absence of control however matters only to this extent on attributionist accounts. In fact, these philosophers claim that in determining a person’s responsibility for her beliefs “there does not seem to be any need to appeal to the agent’s choices, because the thing in question can be seen as directly expressive of her judgments” (Smith 2008: 368). Our beliefs are directly indicative of our judgments in the sense that if we make certain judgments such as healthcare is important and the poor should be afforded the same opportunities as the rich then we will hold particular beliefs such as healthcare should be made accessible for all members of the population regardless of economic status (Levy 4).

Most recently, Angela Smith has defended an attributionist account of moral responsibility on which agents can be held morally responsible for their beliefs. Smith very quickly summarizes her account in the following sentences:

The view, which I have elsewhere called the rational relations view, makes rational judgment rather than choice or voluntary control the basic condition of moral responsibility. To say that an agent is morally responsible for some thing, on this view, is to say that thing reflects her rational judgment in a way that makes it appropriate to ask her to defend or justify it (369).

To reiterate, Smith tells us that our responsibility for things is to be accounted for in terms of our rational activity (Smith 2008: 381). When we morally assess someone for holding a morally objectionable belief or carrying out a morally objectionable action, Smith explains, “we
do not seem to be responding to certain facts about its possible origins in the (agent’s) choices” (Smith 2008: 383). Rather, she states that we seem to be responding to certain judgments that we take to be implicit in her acts or beliefs (ibid). When we criticize a person for example for acting selfishly, Smith seems to be saying that we are responding to certain bad judgments that she has made namely that she has failed to give weight or significance to the interests and needs of others (Smith 2008: 380-381).

Moreover, Smith explains that we seem to be asking an agent to explain or justify her rational judgments and to acknowledge fault if an adequate defense cannot be provided (Smith 2008: 381). Since the justificatory demand implicit in moral criticism is directed at our judgments, Smith argues that the condition of control should be replaced by a “judgment dependence condition” (Smith 2008: 368). In other words, Smith seems to be saying that all that is required for holding someone responsible for some thing is whether that thing reflects her judgments (ibid).

For purposes of clarification, Smith explains that these judgments need not be consciously recognized by the individuals who hold them (Smith 2005: 252). For example, we might not realize until we are faced with the task of sorting through job applications that we tend to judge that minorities are inefficient or unreliable workers. Although we might not have consciously entertained these judgments, Smith explains that these judgments are still attributable to us in virtue of the fact that they are manifested in our actions and beliefs (Smith 2005: 252). According to Smith, we therefore can appropriately be asked to reassess those judgments and to explain, justify, or modify the actions and beliefs that are reflective of those judgments (Smith 2008: 386).
Additionally, Smith tell us that the fact that a person’s judgments might be shaped in various ways by her early attachments and environment does not make it inappropriate to ask her justify or explain those judgments (Smith 2005: 267). She explains that it is not inappropriate because these judgments are still genuinely her judgments about things (ibid). Smith goes on to say that it does not matter how a person’s judgments were formed since all that matters on her account of responsibility is whether the belief or act in question can reasonably be taken to reflect an agent’s judgments (ibid).

Now we might seem unsure about whether the sort of attributionist account of moral responsibility that Angela Smith offers is correct. Initially, we started out with the assumption that an individual can be held morally responsible for some thing just in the case that she has the ability to do some thing at will and the ability to refrain from doing something at will. Now, the attributionists seem to be saying that we can be responsible for our beliefs even though we are not able to choose our beliefs at will. Therefore, I would now like to move on to evaluate Smith’s attributionist account of responsibility in hopes of getting closer to answering the question of whether we can be held morally responsible for our beliefs.

IV.

Before critiquing Smith’s account, let me say a few words about my own views regarding the question of whether individuals can be held morally responsible for their beliefs. My own views fall somewhere in the middle between the attributionist account of moral responsibility and the accounts of responsibility that tend to be supported by individuals who take the “at will” sort of control to be a necessary condition of responsibility (see for example Levy 1-16). To be more explicit, I think that we do not need to control our beliefs in the “at will” sense for us to be held responsible for our beliefs. Yet, I think that the attributionist accounts are under-described.
We are not responsible for our beliefs merely because they are expressive of certain judgments that are attributable to us. Rather, I contend that we are appropriately responsible for our beliefs because in making the judgments that underlie our beliefs we exercise a robust form of agency or control even if there is a sense in which our beliefs are not voluntary. In fact, we will see that the features that render believing less than voluntary are additionally present for intention (Hieronymi 2006: 46).

To begin, Smith seems to offer a thoughtful reply to those individuals who tend to think that we cannot be held responsible for our beliefs if they are not under our control in the “at will” sense. Being control freaks, we tend to want to insert control in all aspects of our ascriptions of moral responsibility. Yet, as Smith points out, in morally assessing someone for things we often are not questioning whether she has chosen to do those things at will.13 Rather, as Smith insists, we often tend to be holding her responsible for judgments that we take to be implicit in the thing in question and to reveal how she regards others. Since often we are not asking whether a person has chosen the thing in question at will, I do not think that it is necessary for us to be able to control our beliefs in the “at will” sense in order to be responsible for our beliefs.

This reply taken from Smith however might not yet be helpful. To elaborate, we end up saying that we can be appropriately held responsible for our beliefs because they are expressive of our judgments. For Smith’s view to be convincing, we need to understand why being expressive of our judgments makes us responsible for some thing. Recall Smith thinks that we can be held responsible for an act or attitude if it is expressive of our judgments even if our judgments are not consciously made because these judgments are manifested in the act or

---

13 I use the word ‘often’ because it may sometimes make sense to ask whether an individual has chosen to do some thing at will. As previously discussed, the problem is that we might for example accidentally knock into some thing, step on some thing, and so on. So learning whether an individual has chosen to do some thing at will is helpful when it comes to our basic bodily movements because this information will help us to determine whether our bodily movements are reflective of our judgments.
attitude in question and in this sense they are attributable to us. But, this seems to just beg the question that we are trying to answer.

I propose that Smith fails to convincingly explain why being expressive of our judgments makes us appropriately responsible for our beliefs because in her zeal to defend her non-voluntarist account of responsibility she overlooks the fact that we exercise a distinct type of control over our beliefs in terms of the judgments that we make. To elaborate, while we might not be able to control our beliefs in the “at will” sort of way, this is not to say that we completely lack control over our beliefs. We are still in control because the move from uncertainty as to whether \( p \) to believing that \( p \) is marked by our settling for ourselves the question whether \( p \). However, recall we cannot settle the question whether \( p \) for any considerations that we take to show in favor of believing \( p \) (Hieronymi 2006: 52). If we for example find reasons that we take to show believing worthwhile or good to do, we have not found reasons that bear on the question of whether \( p \) (ibid).\(^{14}\) Rather, we have found reasons that bear on the question whether believing \( p \) would be good to do (ibid).

Moreover, I am also not saying here that in coming to believe things we make an explicit decision to believe some thing based on our evidence. Matthias Steup defends such an account. Steup proposes, “if by engaging in practical deliberation we can decide what to do, should not we by weighing the evidence, that is, by engaging in epistemic deliberation, be able to decide what to believe?” (Steup 33). He then comes to the conclusion that individuals execute a decision to \( x \) if they \( x \) because they have concluded that their reasons or evidence supports “xing” (Steup 35). Steup goes on to say that the role of a decision in the case of believing is to take us from what he calls “the verdict belief” to what he calls “the object belief” (Steup 37). He

\(^{14}\) These reasons that show believing good to do could include things like the possibility of winning money or being accepted by one’s peers if one believes that \( p \).
explains that the verdict belief is the belief that our evidence supports believing that \( p \); whereas, the object belief is the belief that \( p \) (Steup 36).

I do not agree with Steup that we make an explicit decision to believe in settling for ourselves the question whether \( p \). If all beliefs require such prior beliefs, Richard Feldman points out that it seems as if we must form an infinite regress of ever more complex beliefs in order to form one belief voluntarily (Feldman 2002). In other words, if all our beliefs are acquired by making a decision, then our verdict beliefs must themselves be the product of a decision to believe. Suppose that I see a dog in my neighbor’s yard. I would form the verdict belief that my evidence supports that there is a dog in my neighbor’s yard. However, my verdict belief about the dog would itself have to be the product of some other verdict belief. Furthermore, this higher order verdict belief must itself be a product of a decision to believe and so on.

On my view, in coming to believe that \( p \) we are making judgments in the sense that we are evaluating the evidence and reasons for believing \( p \) and in doing so we are moved to believe \( p \). To elaborate, in coming to believe that \( p \) we actively reflect upon the evidence for and against \( p \). Based on this sort of evaluation, we ourselves come to settle the question of whether \( p \).\(^{15}\) We exercise control in settling the question of whether \( p \) because we judge whether there are good and sufficient epistemic reasons for believing \( p \).\(^{16}\) Also it is important to note that in making such judgments about whether there are good and sufficient epistemic reasons for believing \( p \) we think

\(^{15}\) Incompatibilists sometimes describe a scientist who is capable of manipulating our brain to say produce a specific belief or action. So I want to clarify that we must be the ones actually judging settling the question of whether \( p \).

\(^{16}\) I describe these reasons as “epistemic” to differentiate these reasons from other reasons that we might have for believing things. We might say have pragmatic reason for believing things. Pragmatic reasons for believing refers to those reasons that we take to show believing good to do so, that is, those reasons that settle the question of whether believing \( p \) would be good to do so rather than whether \( p \). But, by epistemic reasons, I mean those reasons that bear on the question of whether \( p \).
about what is true and important to us in the world (Hieronymi 2008: 370). By this I just mean that there is a crucial element involved in belief formation that the evidentialist dismisses in focusing exclusively on the importance of evidence.

All of this is not to say that we can believe whatever we fancy. In other words, my contention is not that we start out with the goal of believing that $p$ and then arrive at the belief that $p$. Rather, my point is that we come to believe $p$ if we judge there to be convincing reasons that show that $p$ is true and represents reality. And we abandon our belief that $p$ if we come to judge after being exposed to certain evidence or reasons that there are not actually sufficient reasons to believe $p$. As an attempt to clarify, let me resort to an example.

Consider my belief that there is a car on the road in front of my house. In coming to believe that there is a car on the road in front of my house, I might judge that there are good and sufficient epistemic reasons for believing that there is a car on the road in front of my house. These epistemic reasons might include that I see a car on the road in front of my house, that I am not too far away from the road to see a car on the road, that there is nothing wrong with my vision, and that cars regularly travel on the road in front of my house. Once I have judged that there are good and sufficient epistemic reasons to believe that there is a car on the road in front of my house, I might then come to believe that there is a car on the road in front of my house.

When the evidence is inconclusive or ambiguous, what we believe may be even more straightforwardly under our control. To elaborate, when the evidence is inconclusive, we judge whether to attend to pertinent moral or logical norms, whether the proposition in question is consistent with other beliefs that we hold, whether to weigh certain pieces of evidence more highly than others, and so on (Govier 654-666). Consider this scenario. After receiving laser eye treatment, suppose that I begin to see red dots on everything in my hospital room. Also suppose
that my doctor has informed me that the treatment that I have received might alter my visual perception and that there were no red dots on anything in my room prior to my treatment. Under these circumstances, I might not weigh too highly the fact that I see red dots on everything in my hospital room. I might even judge that there are sufficient epistemic reasons for believing that there are not red dots on everything in my hospital room namely, the fact that I have undergone on a procedure that could alter my visual perception and that there were not red dots in the room prior to my treatment. I might then come to believe that there are not red spots on everything in the room. If however I was a less logical person, I might have judged differently. I might have failed to reasonably deduce that I had perceived red dots on everything as a result of my eye treatment. And I might therefore have given more weight to the appearance of red dots on everything. I might even have judged that there were good and sufficient epistemic reasons for believing that there were red dots on everything in my hospital room and so I might have come to believe that there were red dots on everything.

Similarly, our beliefs concerning more morally relevant matters seem to be under our control in a distinct sort of way. In coming to believe that the death penalty should be abolished, an individual might judge that there are good and sufficient epistemic reasons for believing that the death penalty should be abolished. These epistemic reasons might include that people could possibly be reformed in prison, that people sentenced to death can sometimes be wrongly convicted of their crimes, that killing of other human beings is always immoral, and that the legal fees accrued by the state for putting a person to death are greater than the costs of keeping a person in prison for life. Once she has judged that there are good and sufficient epistemic reasons to believe that the death penalty should be abolished, she might then come to believe that the death penalty should be abolished. Note different judgments though could have been made.
If the individual was a less compassionate person, she might not have found these sorts of reasons to be sufficient for settling the question of whether the death penalty should be abolished.

Still individuals, who are more sympathetic to the “at will” account of control, might argue that this sort of ability to evaluate and respond to reasons does not amount to the sort of control that is necessary for ascriptions of responsibility. I however think that this objection is not correct. As Sharon Ryan and others have pointed out, the scope of things that can be done “at will” is much narrower than the scope of things that we take to be the paradigm for which we can be held responsible (Hieronymi 2008: 368-369, Ryan 64-66, and Steup 35). That is, we take our intentional actions to be the paradigm for which we are held responsible. But, the problem is that intending does not always seem to be under our control in the “at will” sense.

Consider that intending to act is a matter of having settled for oneself the question of whether to act (Hieronymi 2006: 56). As Pamela Hieronymi states, intending however may seem more voluntary than believing because often an agent can intend to act for any reason that she takes to show acting good to do including that intending to carry out the action would to be good to do (Hieronymi 2006: 57). If you are angry for example because I have no intention to go to the movies with you, that very fact can be reason enough to go to the movies with you. The fact that the intention would make you happy, a fact that shows intending good to do, can be a reason to go to the movies with you, that is, it shows going to the movies with you good to do.

In some cases, the fact that intending to \( \phi \) would be good to do is not taken to be a reason to \( \phi \). Gregory Kavka’s famous toxin puzzle provides us with one such circumstance. However,
such cases also occur in less odd scenarios (Hieronymi 2008: 369). Consider this scenario. Suppose I am upset because you do not intend to take me dinner for my birthday and I care less about actually going out dinner than about the fact that you have no such intention. And suppose you would be willing to intend to take me to dinner as long as you do not actually have to take me to dinner. Unfortunately, it does not seem that you could intend to take me to dinner for my birthday. Recall you cannot intend to take me to dinner because intending is a matter of committing to carry out the action that you intend to carry out (Hieronymi 2008: 369). Since you are not willing to actually take me to dinner, you cannot just intend to take me to dinner because you think intending to take me to dinner would be good to do.

All of this seems very similar to our believing in the sense that we cannot believe $p$ for reasons that we do not take to show that $p$ is true. For example, say that you offered me a million dollars to believe that the capital of the United States is now Las Vegas. Although I might want to believe the capital is now Las Vegas, I would not be able to just believe that the capital of the United States is now Las Vegas. I would not be able to just believe this even though I might want to win the money because winning money for believing the capital is now Las Vegas does not show that Las Vegas is actually now the capital of the United States. Rather, the possibility of winning shows that believing that Las Vegas is now the capital of the United States would be good to do.

17 Kavka describes an eccentric billionaire who has access to technology that allows her to discern other people’s intentions with almost flawless accuracy. She offers to give you a large sum of money on Tuesday if you form the intention on Monday to drink a disgusting but not deadly toxin on Wednesday. To get the money, you do not actually have to drink the toxin you just have to intend to drink the toxin. Michael Bratman explains that the problem is that when Wednesday arrives you will already either have the money or you will not (Bratman 62-63). So you will not have a good reason to drink the toxin but you will also have good reason not to drink the toxin given that it is disgusting. And since intending is a matter of settling whether acting is worthwhile to do, you will not be able to intend to drink the toxin.
I however would be able to believe that the capital of the United States is Las Vegas provided the right kind of reasons namely, good evidence. If for example I read in the New York Times and I see a report on ABC news describing that the capital of the United States is now Las Vegas, I might be surprised but I would believe that the capital of the United States is now Las Vegas. Similarly, you might be able to intend to take me to dinner for my birthday provided the right kind of reasons namely, reasons that show that taking me to dinner good to do. For example, you might be able to intend to take me to dinner if you think that taking me to dinner would be enjoyable. So we cannot just believe p or intend to φ for any consideration that we take to count in favor of doing so since believing represents our take on the truth of p and intending our take on whether acting is worthwhile.

Although our intentions and beliefs are similar in these ways, individuals might object that believing is importantly different than our intentional actions in another way and for this reason we are not in control of our beliefs. When we act, these individuals might argue that we make conscious decisions and plans about how to act (Ryan 67). Consider this example. Right now, I am thinking about going for a walk and taking a break from writing my philosophy paper. I think about whether to go for a walk. I judge that it would be worthwhile to go for a walk because it will re-energize me after spending hours at my computer. So I consciously decide that I will go for a walk and then I go for a walk.

When we are forming our beliefs, these individuals might go on to explain that we do not usually make such conscious decisions to believe things. This seems especially evident concerning out perceptual beliefs. We do not for example say to ourselves that there are tomatoes on the counter so I am going to decide to believe that there are tomatoes on the counter and then do it. Once we see that the evidence supports a proposition, we just believe the
proposition. Yet, I do not think that this comparison shows that we lack control over our beliefs. To elaborate, in actually carrying out certain actions, we often do not exert conscious efforts or make explicit decisions (Ryan 67). Consider walking. After my first few steps, I do not consciously decide to move my limbs, I just move them (Ryan 67). Although I might not consciously decide to move my limbs, they are still under my control because I am the one who is doing the moving (ibid). Similarly, I do not have to explicitly decide or intend to believe something for it to be under my control. Rather, as Ryan explains, I just have to believe what I myself judge to be true (Ryan 67).

Of course we might think that this sort of account of control is an odd way to think of control. To elaborate, we started out with the assumption that a thing is under our control only if we can choose to do or not do at will. And we found this sort of control to be important to us for the kind of responsibility we take to be issue for holding individuals morally responsible for things. We for example pointed out that we normally do not hold each other responsible for physical movements that we do not do voluntarily such as sneezing, seizures, and spasms precisely because we lack the ability to choose at will whether to do them. We also pointed out that we tend to think of control in terms of the way that we can control raising our hands. If we consider that it would be good for us to raise our hands, we can decide to do that, and providing that we are not paralyzed or that someone is not holding down our arms, we can just do it. Similarly, we assumed that we should be able to control our minds in the same way.

Extending this condition of “at will” control to the psychological realm however is a mistake. We do not have the sort of “at will” control over our beliefs precisely because of the type of states that our beliefs are. To elaborate, we do not lack control in the “at will” sense over our beliefs because of any sort of shortcoming in our agency (Hieronymi 2008: 370). We lack
this sort of control over our beliefs because we take our beliefs to represent reality. Nevertheless, there does seem to be some room for control in the doxastic realm. This is to say that our beliefs while not voluntary are not involuntary in the ways that sneezing, spasms, and seizures are. They are not states that simply befall us because believing $p$ is a matter of settling for oneself the question of whether $p$ where that question can be settled if we judge we have sufficient epistemic reasons.

What’s more is that our beliefs are often different than our basic bodily movements because by settling the question of whether $p$ we express our take on what is important and true to us (Hieronymi 2008: 371). Yet, we sometimes move our bodies without settling the question of whether to do so. We could for example loose our footing and move in certain ways that might bring about a bad state of affairs such as a broken vase. So, in determining our responsibility for these movements, we should ask whether the movement was done at will. Otherwise it is difficult to see how these movements could be traced back to us in something other than a causal way.

Likewise, if some state of affairs was not something that you intentionally brought about nor something that was a foreseeable consequence of what you intentionally did, it does not seem appropriate to hold you responsible for that state of affairs in anything other than a causal way (Hieronymi 2008: 372). Suppose for example that I decided to cook you dinner and I asked you beforehand if you had any allergies. Also suppose that you neglected to tell me that you were allergic to peanuts thinking that I would not cook with this ingredient. Consider also that I cooked chicken with peanut sauce and you became violently ill after eating this meal. It does not seem that to morally criticize me for making you ill because your illness was not on my account a foreseeable consequence of cooking dinner for you. Under these types of circumstances, it
therefore seems plausible to say that we cannot be held responsible for certain states of affairs if we did not voluntarily choose to bring them about at will.

To summarize, I intend to strike a middle ground between the attributionist and those who tend to think “at will” control is necessary for moral responsibility. My view can be considered a middle ground between these two accounts because it allows for direct moral responsibility for our beliefs that have not been acquired at will but also includes a control condition for our responsibility for our beliefs.

V.

At this point, I would like to address an objection that might be levied against my account. Since our beliefs track evidence and reasons, individuals might argue that it is only a matter of epistemic luck what sort of judgments we make regarding whether \( p \) in the sense that is only a matter of luck what evidence or reasons we are exposed to. So even though our judgments might violate certain moral norms these individuals might go on to say that we are not fully or robustly morally responsible for our beliefs.

To motivate these individuals’ worries, let us consider a certain racist individual and a certain sexist individual. Suppose an individual grew up and continued to live out his life in a sheltered racist community in which open debate and reflection were discouraged (Wolf 266). As a result, this individual comes to judge that there are good and sufficient reasons for holding certain racist beliefs and so he might come to have certain racial beliefs (ibid). We could tell a similar story about a man that is abused by the important women in his life. Because of this abuse, the man might judge that there are good and sufficient epistemic reasons for believing that women are evil. Consequently, he may come to believe that women are evil.
Some individuals might argue that it seems highly questionable that the racist and the sexist are morally responsible for their beliefs. Given the racist’s background, these individuals might explain the racist does not seem responsible because he had no room for coming to see the reasons or the evidential considerations as to why racism is wrong (Wolf 267). Similarly, individuals might argue that the sexist is not responsible for his beliefs because the abusive environment in which he was raised settled for him the question of whether women are evil.

Although these individuals raise an important worry, these sorts of concerns are not unique to the case of holding individuals responsible for their beliefs. To elaborate, we can no more control what reasons we are exposed to when it comes to our practical deliberation. The reasons that we are exposed to often also seem to be determined by our environments and sometimes by our early attachments. Consider this example. An individual might intend to kill people of a certain religion. He thinks that doing so would be worthwhile because he judges that people of a particular religion are evil and so in killing people of this religion he will prevent an enormous tragedy that they are about to cause. Perhaps, he tends to make such judgments because he grew up in a home reading conspiracy literature about people of this religion.

More generally, my point is that we are responsible for beliefs if we can be responsible for our intentional actions. To elaborate, it is not my intention to defend the concept of moral responsibility. Rather, I am arguing that individuals can be morally responsible for their beliefs if there is such a thing as moral responsibility.18

Conclusion

---

18 I do not however think that these sorts of objections defeat the concept of moral responsibility. As Galen Strawson has pointed out we tend to think that a person need not be responsible for being who he is in order to be held responsible for the things that he does (Strawson 106). In fact, there seems to be two separate questions at issue. These questions are: Is the person responsible for becoming the person who he is? and Is the person responsible for the objectionable attitude or act?
When thinking and writing about moral responsibility, the stakes are high in the sense that an individual will open to moral credit or discredit if he can be found responsible. Given these considerations, we are often very concerned with ensuring that the agent’s act or attitude be traceable back to his own will rather than his circumstances. We therefore tend to think that an agent can be responsible only if he can do some thing at will. Yet, as we saw in this paper, we cannot often carry out actions at will that we take to be the paradigm for which we can be held responsible. So there seems to be good reason to suppose that we need not control things in the “at will” sort of way to be held morally responsible for them. Concerning our beliefs another form of control makes us appropriately responsible for our beliefs. We can be said to be in control of our beliefs because believing $p$ is a matter of settling for oneself the question of whether $p$ where that question can be settled if we judge we have sufficient epistemic reasons to believe that $p$. 
References


