

Autonomy, Agency, and the Value of Enduring Beliefs

(Forthcoming, *Canadian Journal of Philosophy*; please do not cite without permission
Jason Kawall, jkawall@mail.colgate.edu)

There has been much recent interest in questions of value in epistemology (for a thorough overview of this work, see Pritchard 2007), and in what follows I argue that epistemologists concerned with the value of true beliefs and knowledge would do well to devote attention to the enduring nature of beliefs, and in particular to the essential role that they play in constituting agents themselves.

I begin by considering an analogy commonly drawn by epistemologists between acts in the domain of ethics and beliefs in the domain of epistemology, and argue that it is flawed in important respects. I propose that a better, more fruitful analogue for belief would be desire, or a similarly enduring state of an agent. This revision of a commonly used analogy may be of some value in itself, but I further consider how focusing excessively upon the belief-act analogy (with its implicit emphasis on the process of belief-formation) can lead to flaws or shortcomings in our epistemic value theorizing.

Still, this initial work on analogies is intended as a preliminary step, providing an entryway to the more fundamental proposal that we ought to devote greater attention to the enduring nature of beliefs – a nature not captured by the belief-act analogy. I argue that our enduring beliefs help to make us who we are as agents, and it should not be surprising to find significant value to our beliefs arising through this essential, constitutive role. In particular, I argue that enduring, accessible beliefs are crucial to the autonomy, self-control, and authenticity of agents. We would lack effective practical identities without such beliefs.

By seeing (i) that the common belief-act analogy can lead us astray by focusing excessively on the moment of belief-formation (rather than on beliefs as enduring states of agents), and (ii) that agents require enduring, accessible beliefs in order to possess effective practical identities, we are

justified in holding (iii) that epistemologists would do well to devote more attention to the endurance and accessibility of our beliefs when considering questions of epistemic value. Of course one could embrace any of these three theses while rejecting the others – there are not strict entailment relations between them. But they do form a coherent whole – for example, to the extent that we recognize that the endurance and accessibility of our beliefs is crucial to their value, we should not be surprised to find that the belief-act analogy is flawed (precisely insofar as it fails to capture such concerns), and that we might do better to consider a belief-desire analogy. More broadly, an embrace of the positions defended here should encourage an examination of issues often overlooked in discussion of epistemic value.

I

It is common for epistemologists to draw upon work done in ethical theory in developing approaches within the theory of knowledge. Reliabilism (taken broadly) is often treated as being analogous to rule-utilitarianism in the domain of ethics (for example, Firth 1998).¹ There are epistemic deontologists who speak of epistemic duties akin to moral duties, and who draw upon work by deontologists in the moral realm (e.g., Steup 2000, 2001a; Feldman 2001; Russell 2001). And in recent decades, particularly the last several years, there has been much interest in various forms of virtue epistemology that draw upon work in virtue ethics (e.g., Axtell, 2001; DePaul and Zagzebski, 2003; Fairweather and Zagzebski, 2004; Greco 2000; Montmarquet, 1993; Sosa 1991; Zagzebski 1996). In such work beliefs in the epistemic realm are commonly treated as the analogues of actions in the ethical realm. Indeed, it seems that most every epistemologist who draws upon ethics makes appeal to such a comparison.

Jonathan Adler, in *The Ethics of Belief*, compares beliefs to a form of speech act: assertion (Adler 2002; see esp. 274-7). Ernest Sosa likens forming beliefs to an archer shooting at a target, where knowledge is (roughly) akin to an archer's skilful shot that hits its target (Sosa 2003). John

Greco similarly uses various baseball analogies – such as hitting homeruns, or making difficult catches – in his discussions of knowledge and how it differs from mere true belief (Greco 2003).

Richard Feldman has explored analogies between objective and subjective justification in epistemology and ethics (Feldman 1988; see also Russell 2001). He considers several understandings of these notions, drawn from a range of authors. For example:

(D1) Action A is objectively ethically justified for S iff action A is in fact the best of all the actions available to S. (Feldman 1988, 407)

(D6) S is *objectively epistemically justified* in believing p iff S has good reasons to believe p. (Feldman 1988, 411)¹

While Feldman argues that there are important disanalogies between the objective and subjective justification of actions in ethics, and the parallel forms of justification of beliefs in epistemology, note that the analogies he considers are drawn between actions and beliefs.

Finally, in recent work Linda Zagzebski treats beliefs as at least analogous to acts. She writes that

the idea that in a state of knowing the agent gets credit for getting the truth suggests that her epistemic state is attached to her in the same way her acts are attached to her. [...] If believing is like acting, we have a model for the way the agent can get credit for the truth of a belief because of features of the belief that derive from the agent. I propose, then, that [...] *Knowing is related to the knower not as product to machine but as act to agent.* (2003a, 16)

Indeed, Zagzebski is drawn to the stronger claim that believing and knowing are forms of acting:

True believing *is* an intellectual act, or at least, it is strongly analogous to an act. Of course, it is an act that is in part generated by a series of previous acts. Nonetheless, it is enough like an act to make a comparison with overt acts illuminating. I propose, then, to modify my definition of knowledge as follows: knowledge is an act of intellectual virtue. (2003b, 153)

Much can be said for this linking of beliefs and actions across epistemology and ethics. It might be especially useful when we are considering belief-formation, instances of which appear to bear certain similarities to performances of actions. And it is also plausible to think that differences in how beliefs are formed will play some role in answering what Zagzebski has dubbed ‘the value problem’: the question of why knowledge is more valuable than mere true belief.²

Still, for all this, epistemologists may well be focusing too much on processes of belief-formation, and overlooking the significance and value of beliefs as ongoing, enduring states of an agent. There are important disanalogies between beliefs and acts - ones relevant to assessing the value of knowledge and true belief. By moving away from an exclusive focus on the act-belief analogy, and instead comparing beliefs to enduring states like desires, new possibilities with respect to valuation become apparent.

II

Consider the following disanalogies between acts and beliefs:

1. Beliefs and knowledge are partially constitutive of the agent herself – they help to make her who she is, particularly in the case of long-enduring, often-used beliefs. In this, our beliefs are like our fundamental desires in making us the people that we are. But actions are typically occasional, even momentary events; they are not similarly constitutive of the agent. Epistemologists who draw analogies with actions in ethics focus on belief or knowledge-*formation*, but we need to remember that knowledge and beliefs typically endure within us, often long after their formation. This is not true of actions – our actions might reflect or express our psychology, have consequences for it, and so on, but the actions themselves are not part of us as agents.

Still, Zagzebski has suggested that acts are part of agents:

If I am right that knowing is not an output of the agent, it must be a state of the agent. [...] if we think of a belief as part of the agent, the belief can get evaluative features of the agent

in the same way that acts get evaluative properties from the agent. [...] An act is not a product of an agent but is a part of the agent, and the agent gets credit or discredit for an act because of features of the agent. (2003a, 16)

Perhaps the best case for an act being part of an agent would involve a mental act; suppose that we treat belief-formation as a form of mental act. If you ponder long and hard, and eventually arrive at the belief that $2+2=4$, this activity has occurred within you, within your psychology. And to this extent, it may appear to be a constitutive part of you. But compare a word-processing computer program where, when the 'Enter' key is pressed, the program begins a new paragraph in a document. Intuitively, when the program starts a new paragraph this activity is not part of the program; rather it is what the program is now doing. It is an act performed by the program, and not the program itself, even if this occurs 'within' the program, just as mental acts are performed within an agent herself (putting aside issues of externalist mental content).³

Or again, that digestion occurs within a digestive system does not yet show that digestion is a constituent of the system itself. Suppose that a particular instance of digestion changes the digestive system (perhaps an excess of acid produced by the stomach damages the oesophagus); this is akin to a mental act changing the psychology of an agent herself. But even here, the act (the digestion) is a doing which is performed by the system, not the system itself; and this even if the act has an impact upon the system itself. More broadly, an agent's acts are intuitively her 'doings'; they are *performed by* her but they are not part of the agent herself in the same way that her desires, fears, hopes – or beliefs – are part of her.⁴

We might press further by modifying the analogy somewhat; perhaps beliefs are akin to extended actions. For example, suppose Claire is reading *War and Peace*. Of course, she is not occurrently reading this book at every second, but we can understand her reading of the novel as an extended event that takes place over weeks. Such extended actions might seem more analogous to

extended, enduring beliefs (and if so, the act-belief analogy might be preserved). Notice that we have made some progress already. Epistemologists who have drawn upon the act-belief analogy – including in discussions of the value of knowledge and true belief – typically do not focus on such extended acts; rather, they draw on more episodic actions (such as hitting a homerun, or an espresso maker’s making a good cup of espresso). If the analogy can be made to hold, it would still be rather different from what most epistemologists work with.

But the revised analogy still breaks down, in any event. Consider: when Claire is reading *War and Peace* (as an extended action taking place over weeks or months), this seems best understood as consisting in her frequently performing more episodic acts of reading the novel over a period of time. But to say that Claire believes that Paris is in France (as an enduring state) does not similarly suggest that Claire is frequently performing more episodic acts of forming or recalling an occurrent belief over a period of time. Take a 3-month time period. If Claire never picks up the book, it seems false to say that she is reading *War and Peace* during this time. On the other hand, Claire would continue to believe that Paris is in France, even if she never occurrently considers this belief during the same 3-month period. Relatedly, in the action case, the overall extended action seems to be constituted by a series of shorter actions. A war might be understood as an extended event, but one made of up several shorter events (individual battles, troop movements, and so on). But an enduring belief that Paris is in France is not similarly constituted by a series of shorter actions, occurrent beliefs, or instantiations of a belief.

As a further attempt to defend the act-belief analogy, one could argue that in the case of supposedly enduring beliefs we in fact instead have dispositions to form occurrent beliefs under certain conditions - and that the formation of these occurrent beliefs is analogous to an action. For example, suppose that Claire learns that greyhounds can run at speeds up to 45mph, but does not bring this fact to mind for several months. Rather than saying that during this time that she believes

that greyhounds can run at up to 45mph, the objection would instead hold that Claire now simply has a disposition to form, under certain conditions, the occurrent belief that greyhounds can run at up to 45mph (for example, if someone asks her to name a fast-moving dog, or asks for facts about greyhounds). The formation of such occurrent beliefs, according to the objection, would be analogous to the performing of an action.

In responding, we can follow Robert Audi in distinguishing between dispositional beliefs and dispositions to form beliefs (Audi 1994). Compare Fred and Barney. Fred is given a quiz where he must compute $315+22$; he arrives at 337. The next day he is given the same quiz, recognizes the questions, and this time simply remembers his answer from the previous day, 337. Barney does not take the quiz the first day, but on the second day he computes that $315+22 = 337$. We can distinguish between Fred and Barney's performances on the quiz on the second day. Both are disposed to answer "337", but Fred is exhibiting a dispositional belief, while Barney is exhibiting a disposition to belief. As Audi explains,

whereas a belief is – at least in good part – a (state of) readiness to act in certain ways appropriate to its content, at least by affirming the proposition believed, a disposition to believe is a readiness to form a belief. (1994, 423-4)

Fred is disposed to answer "337" because he has a belief that $315+22= 337$, a belief that he formed the previous day and remembers. Barney, on the other hand, is simply good at mathematics, and so when faced with computing $315+22$, he is disposed to form the (true) belief that the answer is 337. This is a new belief for Barney; in the case of Fred we have a dispositional belief that is made occurrent. Thus, while the objection is correct that we can be disposed to form occurrent beliefs, this is not an apt description of what happens in the case of an enduring belief. Note further that on the proposal grounding the objection, Fred would not know his own name, address, and so on,

except when consciously or occurrently considering these items. Presumably this gives us good reason to reject the objection.

2. There are significant relevant differences in voluntariness between beliefs and acts; beliefs seem more akin to desires in this respect. We have many beliefs and desires that form nonvoluntarily – for example, my thirst (and desire for something to drink), or my belief that I am sitting in front of a computer screen as I write this. But we also have desires and beliefs over which some may claim we have greater voluntary control. For example, we might have a general project of leading an environmentally sound life. By reading reliable sources on the topic, we can come to learn about methods for doing so, and come to desire to act in accordance with the proposals we find. Our specific desires here would thus arise through our choosing (voluntarily) to reflect on how to achieve our goal and the information we choose to acquire; there is thus an indirect control exerted over our resultant desires. We might similarly engage in research and come to form beliefs about a range of environmental issues and the plausibility of various stances towards them; our choices of what to read, what evidence to seek out, and so on, will shape our resultant beliefs. The key point here is that desires and beliefs seem to fall across similar ranges with respect to voluntariness, and both seem quite different from actions in this regard. We can take certain steps that may shape our desires or beliefs as in the above example, but this seems an indirect form of voluntariness – we have some control over how we conduct our inquiries, what influences we seek out, etc. But this is not yet to have direct voluntary control over the beliefs and desires that we form; we cannot choose to believe or desire in the same way that we can choose to act. (See Morton 2006 for a discussion of ways in which belief and decision-forming strategies can shape each other).

Acts may also have degrees of voluntariness, ranging from reflex reactions (ducking when you detect an object coming rapidly towards you in your peripheral vision), to paradigmatically voluntary actions (walking to a restaurant you have chosen, where you have carefully deliberated on

the merits of various alternatives). But while this is so, paradigmatic instances of beliefs and desires tend towards the nonvoluntary, whereas paradigmatic acts tend towards the voluntary. Beliefs and desires seem more similar in terms of the forms of indirect control that we can exert over them.

3. Beliefs are unlike actions in having varying degrees of strength. In this beliefs are once again more akin to desires. We can have a strong desire for something to drink, or a weak desire to see a new movie. We are much more prone to try to act on our strong desires. Similarly, we typically have many strongly-held beliefs (perhaps that $2+2=4$) and others that are quite weak (perhaps that Sofia is the capital city of Bulgaria); we will be much more prone to act on our strongly-held beliefs. This can be brought out, for example, by gambling behaviour. On the other hand, acts do not possess degrees of strength – we do not have strong actions for x (or that x), or weak actions for y (or that y). We might be tempted to think that half-hearted or tentative actions might be equivalent to weak desires or beliefs; but notice that when we engage in half-hearted actions it is precisely because we only have a weak desire to perform the action, or only a weak belief that the action will be successful. That is, tentative actions are derivative from tentative beliefs or desires.

We now have a basic case for holding that there are significant, relevant disanalogies between acts and beliefs, and that we may do better to instead treat beliefs as akin to desires, or perhaps similar enduring, constitutive states of agents. Of course there may be certain disanalogies and differences between beliefs and desires; there is no need to claim that they are entirely parallel cases.⁵ What matters is that beliefs and desires are more strongly analogous than beliefs and acts for present purposes. And with this in hand we can turn in the following section to a consideration of how thinking of beliefs as akin to acts, and focusing primarily on questions of belief-formation, might lead to flaws and shortcomings in our theorizing concerning epistemic value.

III

Suppose that we think of beliefs as akin to acts, with knowledge akin to an archer skilfully hitting her target. We can applaud the archer, and be on our way. But consider an all-too common occurrence. The students in an introductory philosophy class are highly attentive (perhaps this is a stretch?), and during the course of a given lecture come to know that affirming the consequent is, in fact, a logical fallacy. They form justified, reliably-produced true beliefs on the basis of the professor's testimony or instruction. Indeed, many might have a period of insight where they truly grasp why it is a fallacy.

So far, so good. And on the act-analogy, it seems we have complete success: after all, we have acts of knowledge-formation. But now comes the all-too familiar: on a quiz only two days later, several students have forgotten most everything from the lecture, and no longer remember that affirming the consequent is a fallacy. Surely there is something going importantly wrong here from an epistemic point of view. The failure of the students' beliefs and knowledge to endure is problematic. But if we are concerned only with the immediate act of belief-formation, it is not clear on what basis we ought to complain. The students did, after all, hit the target.

Suppose you try to learn directions to a friend's house. If you merely understand the directions you are given, and gain knowledge based on the friend's testimony, this is not yet a complete success - you might forget them a few minutes later. You want to remember them at least until you make your trip to the house. You want the belief to endure; you do not merely want a momentary flash of knowledge. Cases like this are common enough - it happens to many of us when we are introduced to several people at some reception or other; we will have knowledge of a new acquaintance's name for a few seconds after being introduced, and then forget. This does not seem to be a complete epistemic success, but again it is hard to locate the problem if we focus only on the process of belief formation.

Returning to our students, what if they were to go back after the quiz, and re-learn the knowledge they once had? If we focus only on individual instances of belief-formation treated as acts, it seems that they should again get equal credit for their later belief formation. Archers can shoot exactly the same target over and over, and gain credit each time. Compare another case: suppose we have a person who suffers from a form of amnesia whereby she cannot add new beliefs to her long-term memory. Every few hours she relearns the same piece of information. Is she an epistemic exemplar? On the basic belief-as-act model, it seems she might be; after all, she keeps performing individual acts of intellectual virtue. Each time she successfully gains knowledge. But presumably she is hardly an epistemic exemplar – she merely relearns the same item of information over and over.

Perhaps we could find grounds for holding that, even on the act-model of knowledge, we cannot gain epistemic credit by simply forgetting and relearning the same items of knowledge over and over. We can leave this as an open question. But even if such a restriction could be justified, the underlying problem remains. Imagine an agent who gains knowledge upon knowledge – she never learns the same proposition twice. But again, she only retains any given belief for five seconds, and then forgets it. Surely we still do not have an exemplary epistemic agent, even if she is akin to an archer who always, skilfully, hits a wide-range of targets.⁶ The difference between recognizing beliefs as enduring states of an agent, and treating them as a mere collection of isolated actions is thus significant when assessing the epistemic performance of an agent.

IV

To this point then, we have considered (i) ways in which beliefs are disanalogous to acts (and instead seem more akin to desires), and (ii) ways in which appeal to the belief-act analogy (and focusing on belief-formation) can lead to flaws or shortcomings in addressing questions of epistemic value. In this section we will turn to a consideration of how treating beliefs as enduring states that

partially constitute agents can help us to explain the value of both true and justified beliefs, with an emphasis on the former.⁷ It will first be shown, in passing, that common accounts of the value of knowledge and true belief typically presuppose enduring beliefs. More significantly (I hope), I will argue that enduring, accessible beliefs are necessary for an agent to possess autonomy, a practical identity, self-control, and authenticity. The crucial role of belief in such states becomes evident when we explicitly shift our focus to beliefs as enduring states of agents with ongoing projects and commitments. Finally, it should be acknowledged that not all philosophers would deem such states as justified or true belief to be valuable (indeed, some will question whether justification is a coherent concept), and that there are other epistemic states that may be of value, such as wisdom or understanding. But for present purposes, we will focus on true and justified beliefs as enduring states that are widely and plausibly held to be valuable.

Why are true beliefs valuable? Some would argue that true beliefs simply are intrinsically valuable as such.⁸ Many epistemologists would stress that true beliefs are often crucial to the successful pursuit of our practical projects, and that we can distinguish between more and less valuable true beliefs, depending on what relation they bear to our interests. Hilary Kornblith, for example, explicitly appeals to the instrumental or pragmatic value of knowledge and true beliefs:

Knowledge is of extraordinary instrumental value, for it allows us to achieve our biologically given goals, as well as our more idiosyncratic individual goals, whatever those goals may be. This makes it a condition that is universally valuable. (2002, 160)

Kornblith emphasizes the importance of true beliefs in making practical assessments of how to achieve our goals. He makes use of an example where we are choosing which of two toasters to purchase:

In choosing between the two toasters [...] we must figure out the consequences of the two purchases; we must assign values to each of them; we must do some arithmetic. If we performed this calculation by using a cognitive system that gave us true beliefs, we would thereby be informed about the actual consequences of purchasing each toaster, what it is we actually value, and the extent to which these consequences actually produce those things we

value. We would thus come to know which toaster better serves our interests, whatever those interests may be. (2002, 155)

If we were to use false inputs, and / or make use of cognitive systems that produce false outputs, then our assessment of which toaster we ought to purchase would likely be skewed. More broadly, poor calculations of this kind would lead us to be thwarted in attempting to pursue our interests.

Finally, Kornblith suggests that

My argument for the importance of truth turned on its being implicated in certain cost-benefit calculations; the calculations that we need to perform in making choices among alternative courses of action must be done accurately. But various goals other than truth are likely to be implicated in this task as well. A system of evaluation that was perfectly accurate but could not perform its evaluations in real time would be of little value. (2002, 158-9)

Kornblith could well be correct in his claims; and he is careful to stress that truth may not be our only goal in belief-formation. Still, notice that we need our true beliefs to last long enough that we can act on the results of our assessments. If we cannot access them, or if they fail to endure, our true beliefs will fail to have the instrumental value suggested by Kornblith and others.

Zagzebski argues that knowledge and true belief are important from a moral point of view.

She suggests that

When something of moral importance is at stake when someone performs an act S , then if S is a case of acting on a belief B , it is morally important that B be true. (2003a, 19)

Suppose now that an act S is a case of acting on a belief B and that act S is an instance of an act type that is morally praiseworthy in the right conditions. I propose that act S is credited to the agent only if the truth of belief B is credited to the agent. (2003a, 19-20)

Intuitively, on Zagzebski's proposal, if you were to give the correct medication to a person suffering from a seizure due to a rare illness, this might be a morally praiseworthy action. But if you simply went to a shelf with a wide range of medications, and chose the bottle with the most tempting, candy-like pills, any belief you may have that this medication is the correct one seems quite unjustified. And as such it seems you deserve no moral credit for giving the person suffering the

seizure this medication - you were merely lucky, and could very well have caused even more serious harm. We can agree with Zagzebski on this point, and again need only stress the importance of enduring knowledge here; mere momentary acts of knowledge-formation are not enough. Without stocks of relevant, enduring knowledge, we would be unable to act on our beliefs (either as we lack such beliefs entirely – if we were to forget them, or as we lack access to them in circumstances where we need them).

Thus, in order to possess the instrumental and moral values posited by Kornblith, Zagzebski, and others, beliefs need to be accessible, and to endure long enough to be used; this can be overlooked if we focus on belief-formation (and the act analogy).⁹ More importantly, we can go beyond such practical and moral concerns when considering the value of knowledge and true belief. Enduring beliefs are essential to our personal identities. As noted earlier, our beliefs are part of our constitution, part of who we are as agents; we can now treat this more fully.

First, consider broadly psychological accounts of personal identity such as that of Derek Parfit (1984, part III). Roughly, Parfit treats personal identity as a matter of degree, varying with psychological continuity and connectedness (“Relation R”).¹⁰ Connectedness consists in the common psychological connections that we find in people in ordinary life – such things as memory of past events, persisting beliefs, desires, and goals, similarities in character or temperament, and so on. Continuity is a matter of having overlapping chains of (high levels) of such connectedness. Obviously, on most any psychological account of personal identity, connectedness and continuity of belief will be of great importance. If there were a complete absence of connectedness in beliefs between two persons over time (i.e., they do not share any memories / beliefs about what they have done, what they value, who are their friends, any items of general knowledge, etc.), we would be hard-pressed to see them as sufficiently similar to see the first person as surviving in the later person.

We can also consider the importance of beliefs to personal identity where this is taken not in a metaphysical sense, but as a matter of practical identity. What do we value and stand for? What makes us the persons that we are? Cox, LaCaze, and Levine, in discussing accounts of personal integrity, draw attention to work by Bernard Williams:

One option here is to define integrity in terms of the commitments that people identify with most deeply, as constituting what they consider their life is fundamentally about. Commitments of this kind are called ‘identity-conferring commitments’ or sometimes ‘ground projects’. This view of integrity, the identity view, is associated most closely with Bernard Williams [...] The idea is that for people to abandon an identity-conferring commitment is for them to lose grip on what gives their life its identity, or individual character. An identity-conferring commitment, according to Williams, is ‘the condition of my existence, in the sense that unless I am propelled forward by the conatus of desire, project and interest, it is unclear why I should go on at all.’ (2005; they cite Williams 1981a, 12)

On the identity view of integrity, *to act with integrity is just to act in a way that accurately reflects your sense of who you are*; to act from motives, interests and commitments that are most deeply your own. (2005, emphasis added; they cite Williams 1981b, 49)

Without enduring, accessible beliefs about what we value and the content of our projects, it is unclear that we could even have a sense of who we are; if so, it seems we would be unable to act with integrity on Williams’ account. More broadly, without forming and maintaining accessible beliefs about our values and projects, how these values might be best pursued, how they might be best balanced, and so forth, it seems that we would be unable to effectively pursue our ground projects and other commitments. To the extent that these are crucial to our personal identity in Williams’ sense, it seems we would be incapable of leading a life worth living – we could well lack an effective practical, personal identity.

More broadly yet, even if we do not appeal to Williams’ identity-conferring commitments, our potential for leading a good life would be severely restricted by a lack of relevant, enduring, and accessible beliefs. For example, how could we maintain or develop a friendship if we could not remember who a given person is from day to day (or, as in the case of anterograde amnesia, we

cannot transfer short-term memories to long-term memory, so that we may recall less than 30 seconds at any given time)? It would be as if we were simply meeting a stranger over and over. How could the foundational trust that is a crucial element to strong friendships develop? We would be unable to recall our friend's preferences, projects, abilities, or flaws. Similarly with other projects; if you wish to train to run a marathon, your project will be undermined if you consistently forget about your project when it is time for you to train, or if you cannot access beliefs about what training you require, and so on. If you cannot recall past mistakes, or remember how to improve, you will be unable to learn from these past failings. Notice that we are moving beyond mere difficulties in performing a single action - the sort of pragmatic concern that is typically raised when considering the value of true belief and knowledge. Rather, we are considering the role of belief in leading a good life as a whole.

Alfred Mele brings a related concept to our attention: self-control. Loosely, we might say that self-control is what a person lacks when she suffers from *akrasia*. According to Mele, to act with self-control is, among other things, to act in accordance with one's best judgment, which he explains as follows:

A judgment that one's *A*-ing would be best, in this stipulated sense of "best judgment," is what I will call a *type-B* belief that one's *A*-ing would be best: a *type B-belief* is, by definition, a belief nondeviantly formed as a conclusion of first-person *practical evaluative reasoning*, such reasoning being understood as an inferential process, involving evaluative premises, driven by motivation to settle upon what to do. (1995, 26)

Mele suggests that standard cases of akratic action are those where a person acts contrary to her best judgment, in the sense given above. But what shall we say if an agent forgets her evaluative, 'type-B' beliefs within minutes of forming them, before she can act – or never forms such beliefs in the first place? What if she cannot access her values (and corresponding evaluative premises as she attempts

to reason) or cannot remember relevant descriptive beliefs, or can only access such beliefs when given prompts? In these cases she may be unable to arrive at a best judgment at all as to what to do; and if so, she would be incapable of self-controlled action.

John Christman makes the following proposal concerning personal autonomy:

an agent is autonomous vis-à-vis some desire if the influences and conditions that give rise to the desire were factors that the agent approved of or did not resist, or would not have resisted had she attended to them, and that this judgment was or would have been made in a minimally rational, non-self-deceived manner. (1991, 22)

On this account of autonomy, an agent would need to form and appeal to beliefs (or be capable of doing so) concerning the influences and conditions that gave rise to her desire(s). But if she cannot remember what these beliefs are, or in the course of her evaluation comes to forget her beliefs, or again if she has the beliefs, but simply cannot access or appeal to them while she performs (or would perform) her evaluation of a desire and its antecedent conditions, it seems she would be unable to be autonomous with respect to this desire. In particular, she could not make a minimally rational judgment concerning her desire and the factors that gave rise to it in such circumstances, and as such her autonomy would be undermined.

Similarly, Gerald Dworkin has suggested that

Autonomy is a second-order capacity to reflect critically upon one's first-order preferences and desires, and the ability to either identify with these or change them in light of higher-order preferences and values. By exercising such a capacity we define our nature, give meaning and coherence to our lives, and take responsibility for the kind of people we are. (1988, 108)

Here again, an agent would need to be capable of forming beliefs about her first-order preferences and desires; but beyond this, she would need to have on-going access to them. If she can only

maintain such beliefs for a moment or two before forgetting, or can only access them under highly unusual conditions, she would lack autonomy insofar as she would lack the capacity to reflect critically upon her preferences and desires. An agent who cannot remember her higher-order values and preferences -and thus cannot shape her first-order desires in light of them - hardly seems an autonomous agent.

A further, related condition that would be undermined by a lack of ongoing, accessible beliefs is authenticity. While ‘authenticity’ is, of course, a technical term of art in Sartre and others, we can focus here on a proposal from Joel Feinberg that captures many common intuitions:

A person is authentic to the extent that ... he can and does subject his opinions and tastes to rational scrutiny. He is authentic to the extent that he can and does alter his convictions for reasons of his own, and does this without guilt or anxiety ... he will select his life style to match his temperament, and his political attitudes to fit his ideals and interests. (1986, 33; cited in Mele 1995, 156)

On Feinberg’s account of authenticity, enduring, accessible beliefs will again be of great importance. If an agent cannot remember or recall his opinions (or never forms any), rational scrutiny of them is not possible. Even if the agent can recall such first-order opinions or beliefs, it is unclear that she could subject these or her tastes to rational scrutiny unless she has access to what her values are, her beliefs about how to achieve these, and so on. If she cannot remember the possible lifestyles she may choose from (or fails to form beliefs about them), or cannot form an enduring belief that certain political attitudes cohere with her ideals, and so on, her potential for authenticity (on Feinberg’s account) is severely limited.

Now unlike pragmatic or instrumental value, it may seem that with respect to being authentic, autonomous, self-controlled, and so on, that the truth of our beliefs does not matter. False inputs and outputs concerning which toaster to buy will typically fail us; we will not buy the

better toaster. But we will still have an ongoing personal identity on Parfit's account, for example, even if the memories and beliefs that help to provide psychological connectedness are largely (or even entirely) false. We can still make choices and act upon our beliefs, even if these are false. In responding to this claim, we can focus on the importance of true beliefs to autonomy, and apply our results to authenticity, self-control, and maintaining a practical identity. With respect to personal identity (in the metaphysical sense), we can concede that we will possess such an identity even if our enduring beliefs turn out to be false; still, even here, we can at least point to the importance of enduring beliefs and memories.

Why might true beliefs matter to autonomy and related states? Crucially, if our reasoning about our values and how to achieve them is based fundamentally on false beliefs, this seems likely to interfere with our autonomy; the actions we undertake on the basis of such reasoning would often not fully reflect our values, and would seem to be less than fully autonomous to that extent.¹¹ For example, if you are making a decision about whether to undergo a highly-invasive medical procedure, but you have false beliefs about the risks involved, the seriousness of your condition, the likelihood of success, and so on, then it seems your decision is not in fact fully-autonomous; it does not genuinely reflect your values. If you were provided with accurate information you would come to reject the decision you have arrived at; and this revised decision would seem to be more truly autonomous, more a reflection of yourself and your values. Or consider the humanity as an ends formulation of Kant's Categorical Imperative. Why should we not lie, according to this formulation? A commonsense understanding of Kant's position here would hold, roughly, that we would not be treating the person lied to as a rational, autonomous agent; giving her false data does not allow her to make truly autonomous decisions that reflect her values.

Similarly, if we have false beliefs about our desires, their origins, and so forth, our authenticity can be undermined. If we build our practical identity upon a set of false beliefs about

why we possess the values that we do, or if we form false beliefs about what we desire or value, a certain lack of authenticity could emerge. For example, if you do not recognize that your desire to become a doctor is exclusively a result of indoctrination by your parents, you might be living a lie (as it were) in pushing yourself into medical school, and finding yourself disillusioned and unsatisfied despite your success. If we believe that we value fame while we in fact value and desire to be respected or loved, then again, it seems likely that we will feel unfulfilled, and that there will be a certain lack of authenticity to our life as we pursue fame rather than love. We can imagine individuals who form false beliefs about their sexual orientation, their career or other aspirations, and so on; any of these could produce a lack of authenticity in the individual's practical identity. And again, false empirical beliefs, false beliefs about our systems of values, and so on, would seem to render our 'best judgements' in Mele's sense flawed and not truly our best judgements; if so, our capacity for genuinely self-controlled action would seem to be impaired to the extent that false beliefs figure in our assessments.

We can thus see the importance of true beliefs with respect to these valuable states; to the extent that false beliefs are implicated in our reflections about our values and projects (or we lack relevant information), our assessments and resultant actions will be flawed and not genuinely authentic, autonomous, or self-controlled. We need enduring true beliefs in order to be autonomous, authentic agents – and surely this explains much of the value of such beliefs. We thus answer the question not only of which toaster to buy - but whether we can be autonomous, self-controlled, and authentic in even deciding to buy a toaster. This seems more fundamental.

But what of justification? That is, does it matter (with respect to autonomy, authenticity, and such states) whether our enduring, self-constituting beliefs are epistemically justified? Here we must be more tentative. It could be that for any individual belief, it does not matter whether it qualifies as knowledge (or at least as a justified true belief) or not, for many pragmatic purposes. If you have a

true belief that Paris is in France, you can still correctly answer quiz questions, book a flight, and so on, even if your belief is not justified. On the other hand, even if your false belief that Paris is in Scotland is justified, you will typically fail to answer quiz questions correctly, and you will have trouble booking a flight. The value of justified belief, then, is perhaps better understood in terms of the agent's ongoing behaviour and faculties. For example, Timothy Williamson suggests that knowledge will typically tend to be more enduring than mere true belief insofar as the former is less vulnerable to rational undermining by future evidence.¹² To the extent that this is so, an agent's autonomy, authenticity, and so forth would also tend to be more resilient.

Justified beliefs may also suggest that the agent is likely to form true beliefs in appropriate circumstances in the future; her justified belief indicates that she can – and at least sometimes does – makes use of truth-conducive belief-forming mechanisms. And to the extent this is so, there is a greater likelihood of her maintaining her autonomy, authenticity, and so on in the future (by being less likely to form false beliefs that would undermine such states). We can hold that an agent with a justified true belief is a greater success than an agent with only a true belief insofar as she is a better epistemic agent than the latter agent – she makes use of better belief-forming mechanisms, ones that should aid her in the future, regardless of the fact that both agents have beliefs with the same content in this particular case.¹³ Again, we need to focus on a longer-term perspective, looking at the enduring, ongoing character and constitution of the agent, rather than just the immediate belief and its formation.¹⁴

V

In assessing the value of possessing true and justified beliefs we should not only consider whether (for example) an agent can successfully choose a better toaster – or satisfy other pragmatic concerns, even if these are important in their own right, and require enduring beliefs. At least equally important is the question of whether an agent can be autonomous, authentic, or self-

controlled in choosing to buy a toaster (or indeed, in her life as a whole), without having enduring, accessible stocks of relevant true beliefs - beliefs about her values, her needs, what toasters and other objects do, what alternatives she has, and so on. I have argued that this is not possible, and that we thus find a still more fundamental value to our enduring, accessible true and justified beliefs. In order to possess such valuable states as authenticity, self-control, or autonomy, an agent requires accessible, enduring true beliefs.

More broadly, given (i) the importance of enduring and accessible beliefs to autonomy, self-control, etc., and (ii) the shortcomings in the belief-act analogy (and its focus on belief-formation), we have justification for holding (iii) that epistemologists would do well to explicitly acknowledge the endurance and accessibility of beliefs in addressing questions of epistemic value. Acquiring knowledge is often treated as the primary goal of an epistemic agent - and certainly we need not dismiss the value of knowledge as such. But clearly, even if an agent acquires knowledge in a given case, this is not yet sufficient to be a complete success from an epistemic point of view. An agent could forget the newly-acquired knowledge moments later, or only be able to access the knowledge under certain highly-constrained conditions. An agent who acquires knowledge but soon forgets is presumably not as successful an epistemic agent as one who acquires the same information, but is able to recall it when confronted with a multiple-choice quiz a week later. More successful yet would be an agent who could easily access the same information spontaneously, without need for cues. And what shall we say about an agent, who, when asked to recall a friend's address would (incorrectly) reply "621 Maple St.", but who would recognize her friend's correct address, "612 Maple St." when she sees it written on a sheet with several other addresses?¹⁵ Does she know her friend's address? How should we evaluate the value of the knowledge she may possess in such a case? In evaluating an agent's epistemic success we need to take into account not just the content of her beliefs, and their epistemic connections to her other beliefs (in terms of justification), but also

their status as psychological states – their accessibility and endurance, how they fit into her overall cognitive economy.

We typically hope and expect our beliefs and knowledge to endure and to be accessible to us; as epistemologists we ought to acknowledge these expectations when assessing the success of epistemic agents. Mere acts are comparatively fleeting, and we can be misled in our value theory if we focus too much on the process of belief-formation. Enduring beliefs, in conjunction with our desires, hopes, fears, and other such states, are fundamental to our personal identities, to making us the agents that we are; without beliefs it is not clear that we could even possess a practical identity. Explicitly acknowledging this constitutive and enduring role may well provide us with new insights into epistemic value and what it means to be a successful epistemic agent. In forming and maintaining justified, true beliefs, we are forming and maintaining ourselves as autonomous, authentic, and self-controlled agents.¹⁶

¹ I argue elsewhere that reliabilism, with its emphasis on the agent's own personal accumulation of reliably-produced true beliefs, is better understood as analogous to a form of ethical hedonistic egoism. If we were to take the rule-utilitarianism analogy seriously, we would arrive at an epistemic position where an agent ought to maximize reliably-produced true beliefs, whether held by herself or by others. See Kawall, 2002.

² For Zagzebski's discussion of this value problem, see her (2000; 2003a; 2003b). Duncan Pritchard distinguishes three value problems: the primary (why is knowledge more valuable than mere true belief?), the secondary (why is knowledge more valuable than any proper subset of its constituents?), and the tertiary (why is the value of knowledge of a different kind than any proper subset of its constituents?). See Pritchard (2007).

³ Should the action of beginning a new paragraph be understood instead as a part of the document that is being processed, rather than the program? The shift in analogies does not help; the event of a new paragraph being created is not part of the document itself, even if the new paragraph (the result of the event) *is* part of the document.

⁴ An agent's actions (and their consequences) may thus have important impacts upon her, reshaping the course of her life and character. Still, such actions are not themselves constituents of the agent; they are rather expressions of her character, performed by her, that have further impacts upon her psychology. (See Kane's discussion of self-forming actions, in his 1996, especially 75-8. On the current proposal such actions would perhaps involve a radical *expression* of agency, but would not themselves be constitutive parts of an agent).

⁵ For example, beliefs and desires are often claimed to have different directions of fit (such that, roughly, we typically ought to change our beliefs to fit the world, but change the world to fit our desires). We typically also hope that our desires are satisfied, and thus come to an end, while we hope that our beliefs endure, without ending. But of course parallel disanalogies would also exist with respect to acts and beliefs: we try to change the world when we act, and we typically hope that our actions are successful and will come to an end. Some disanalogies may be inevitable, simply given that beliefs and desires (and other enduring states) *are* different states.

⁶ Might virtue epistemologists claim that if an agent's beliefs are so quickly forgotten, that they could not have arisen from stable, enduring belief-forming dispositions, and so would not qualify as knowledge or even as justified? Such a claim would be difficult to defend; after all, it seems we could have stable dispositions to form true beliefs that endure only for seconds. More broadly, while virtue epistemologists have embraced the importance of enduring belief-forming mechanisms and traits of intellectual character, they have not similarly drawn attention to the importance of

enduring, accessible beliefs. Still, given their emphasis on enduring traits, virtue epistemologies might be especially amenable to similarly acknowledging the importance of enduring beliefs.

⁷ We will not directly consider the question of the value of knowledge over its constituents. Kvanvig (2003) draws attention to difficulties in accounting for such value, and I lack complete responses to these worries at this time. But regardless of this, discussion of justified and true beliefs will assist in understanding the value of knowledge to the extent that they are components of knowledge itself.

⁸ For discussion of this possibility, see Kvanvig (2003, 38-42); see also David (2001).

⁹ A distinction could perhaps be drawn here between memory as a storehouse of enduring beliefs, and memory as an active faculty capable of bringing enduring beliefs to mind (i.e., to bring them to occurrent attention); an agent could presumably possess a strong memory in one of these senses, while having a weak memory in the other. Much more could – and ultimately should – be said about the nature of memory, but such issues, lying in the domain of philosophy of mind, are beyond the scope of the current paper.

¹⁰ More precisely, Parfit suggests that X at t_1 is the same person as Y at t_2 just in case X is uniquely psychologically continuous with Y , i.e., just in case there are overlapping chains of strong psychological connectedness obtaining between X and Y , and the chains obtain *only* between X and Y (and not some other simultaneously-existing Z). See Parfit (1984, 207).

¹¹ See, for example, Aristotle (1985, book iii), or the discussion of deception and an agent's "informational condition" in Mele (1995, 180).

¹² Williamson (2000, 78-80). Williamson would not analyse knowledge in terms of justified belief, but his suggestion plausibly applies equally to justified belief.

¹³ See Goldman and Olsson (forthcoming) for more detailed discussion of such a proposal.

¹⁴ The current suggestion goes beyond looking at the enduring nature of belief, and instead turns to the enduring nature of agents more broadly – their ongoing traits of intellectual character, their belief-forming mechanisms, and so on, when considering the *value* of possessing justified beliefs. This is not to hold that a belief's being justified is best understood in terms of evaluating the mechanisms, virtues, or agent from which it flows. This latter approach, characteristic of virtue epistemology, strikes me as plausible, but is not being defended here.

¹⁵ For a helpful discussion of varieties of 'in-between' believing, see Schwitzgebel (2002, section 4.3).

¹⁶ Earlier versions of this paper have been presented at the University of Glasgow, the University of Aberdeen, and meetings of the Canadian Philosophical Association, American Philosophical Association (Central Division), and the Illinois Philosophical Association. I would like to thank the audiences and my commentators on all of these occasions. I would especially like to thank Peter Baumann, Martijn Blauuw, Michael Brady, Tom Carson, Mylan Engel, Yertle LaTortue, Fiona MacPherson, John Mahlan, Adam Morton, Baron Reed, Zach Silver, and Linda Zagzebski for their helpful comments. Finally, many thanks to two anonymous referees and the editorial board of this journal for several helpful comments and suggestions.

REFERENCES

- Adler, J. 2002. *Beliefs Own Ethics*. Cambridge: A Bradford Book, The MIT Press.
- Aristotle. 1985. *Nicomachean Ethics*. trans. Terence Irwin. Indianapolis: Hackett.
- Audi, R. 1994. "Dispositional Beliefs and Dispositions to Believe." *Noûs* 28: 419-34.
- Axtell, G., ed. 2001. *Knowledge, Belief, and Character*. Lanham: Rowman & Littlefield.
- Christman, J. 1991. "Autonomy and Personal History." *Canadian Journal of Philosophy* 21: 1-24.
- Cox, D., La Caze, M., Levine, M. 2005. "Integrity." in *The Stanford Encyclopedia of Philosophy* (Fall 2005 Edition), E. Zalta, ed. <<http://plato.stanford.edu/archives/fall2005/entries/integrity>>.

-
- David, M. 2001. "Truth as the Epistemic Goal." in *Knowledge, Truth, and Duty: Essays on Epistemic Justification, Responsibility, and Virtue*, M. Steup, ed. New York: Oxford University Press: 151-69.
- Dworkin, G. 1988. *The Theory and Practice of Autonomy*. Cambridge: Cambridge University Press.
- Feldman, R. 1988. "Subjective and Objective Justification in Ethics and Epistemology." *The Monist* 71: 405-19.
- . 2001. "Voluntary Belief and Epistemic Evaluation." in *Knowledge, Truth, and Duty: Essays on Epistemic Justification, Responsibility, and Virtue*, ed. M. Steup. New York: Oxford University Press: 77-92.
- Feinberg, J. 1986. *Harm to Self*. New York: Oxford University Press.
- Firth, R. 1998. "Epistemic Merit, Intrinsic and Instrumental." in *In Defense of Radical Empiricism: Essays and Lectures by Roderick Firth*, ed. J. Troyer. Lanham, MD: Rowman & Littlefield.
- Goldman, A and Olsson, E. forthcoming. "Reliabilism and the Value of Knowledge." in *Epistemic Value*, ed. D. Pritchard, A. Millar, and A. Haddock. New York: Oxford University Press.
- Greco, J. 2000. *Putting Skeptics in Their Place: The Nature of Skeptical Arguments*. New York: Cambridge University Press.
- . 2003. "Knowledge as Credit for True Belief." in *Intellectual Virtue*, ed. M. DePaul and L. Zagzebski. New York: Oxford University Press: 111-34.
- Humberstone, L. 1992. "Direction of Fit." *Mind* 101: 59-83.
- Kane, R. 1996. *The Significance of Free Will*. New York: Oxford University Press.
- Kawall, J. 2002. "Other-Regarding Epistemic Virtues." *Ratio* 15: 257-75.
- Kornblith, H. 2002. *Knowledge and its Place in Nature*. New York: Oxford University Press.

-
- Kvanvig, J. L. 2003. *The Value of Knowledge and the Pursuit of Understanding*. New York: Cambridge University Press.
- Mele, A. 1995. *Autonomous Agents: From Self-Control to Autonomy*. New York: Oxford University Press.
- Montmarquet, J. A. 1993. *Epistemic Virtue and Doxastic Responsibility*. Lanham: Rowman & Littlefield.
- Morton, A. 2006. "Knowing What to Think about: When Epistemology Meets the Theory of Choice." in *Epistemology Futures*, ed. S. Hetherington. New York: Oxford University Press: 111-30.
- Parfit, D. 1984. *Reasons and Persons*. Oxford: Oxford University Press.
- Pritchard, D. 2007. "Recent Work on Epistemic Value." *American Philosophical Quarterly* 44: 85-110.
- Russell, B. 2001. "Epistemic and Moral Duty." in *Knowledge, Truth, and Duty: Essays on Epistemic Justification, Responsibility, and Virtue*, ed. M. Steup. New York: Oxford University Press: 34-48.
- Schwitzgebel, E. 2002. "A Phenomenal, Dispositional Account of Belief." *Noûs* 36, 249-75.
- Sosa, E. 1991. *Knowledge in Perspective: Selected Essays in Epistemology*. New York: Oxford University Press.
- . 2003. "The Place of Truth in Epistemology." in *Intellectual Virtue*, ed. M. DePaul and L. Zagzebski. New York: Oxford University Press: 155-79.
- Steup, M. 2000. "Doxastic Voluntarism and Epistemic Deontology." *Acta Analytica* 15: 25-56.
- . 2001a. "Epistemic Duty, Evidence, and Internality." In *Knowledge, Truth, and Duty: Essays on Epistemic Justification, Responsibility, and Virtue*, ed. M. Steup. New York: Oxford University Press: 134-48.
- Williams, B. 1981a. "Persons, Character, and Morality." in *Moral Luck: Philosophical Papers 1973-1980*, ed. B. Williams. Cambridge: Cambridge University Press.

-
- 1981b. "Utilitarianism and Moral Self-Indulgence." in *Moral Luck: Philosophical Papers 1973-1980*, ed. B. Williams. Cambridge: Cambridge University Press.
- Williamson, T. 2000. *Knowledge and its Limits*. New York, Oxford University Press.
- Zagzebski, L. 1996. *Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundations of Knowledge*. New York: Cambridge University Press.
- . 2000. "From Reliabilism to Virtue Epistemology." in *Knowledge, Belief, and Character*, ed. G. Axtell. Lanham: Rowman & Littlefield: 113-22.
- . 2003a. "The Search for the Source of Epistemic Good." in *Moral & Epistemic Virtues*, ed. M. Brady and D. Pritchard. Malden, Ma: Blackwell: 13-28.
- . 2003b. "Intellectual Motivation and the Good of Truth." in *Intellectual Virtue*, ed. M. DePaul and L. Zagzebski. New York: Oxford University Press: 135-54.