

The Authority of Memory

David Owens

Nothing is more common than for us to continue to believe without rehearsing the reasons which led us to believe in the first place. It is hard to see how it could be otherwise. Were we obliged constantly to re-trace our cognitive steps, to reassure ourselves that we are entitled to our convictions, how could we ever move forward? We have probably forgotten why we adopted many of our current beliefs and even if we could dredge the evidence for them up from memory, we couldn't do this for more than a tiny subset of our beliefs at any one time. Since enquiry involves a reliance on many different beliefs, progress is possible only if we can use established results in future deliberation without re-fighting the battles of the past.¹

But this plausible thought appears to conflict with another, that we should believe only where we have adequate evidence: rational belief must be based on evidence for the proposition believed. Now one might quibble over what exactly 'have evidence' means. Is it required that whenever the belief comes to mind, so too does the evidence on which the belief is based? Or is it sufficient that one be capable of rehearsing this evidence? Either way, human beings are very often unable to satisfy this demand in respect of beliefs on which they happily rely. If this is illicit, enquiry must be reined in, constrained by our memory's inability to retain more than a fraction of the evidence relevant to the beliefs we formed at various points in the past.

Epistemologists have responded to this tension in several ways. Externalists simply drop the demand that belief be based on reasons at all. Internalists try to find evidence on which rational memory belief might be based. There are two internalist strategies here. One claims that certain forms of empirical argument are generally available to underwrite the memory beliefs of the rational person, arguments which one can rehearse even if one can't recall the specific grounds on which one formed the belief in the first place. The other strategy simply asserts that memory beliefs have a *prima facie* authority, that one is entitled to rely on them without any justification, provided one has no grounds for doubting them.

I agree with the internalist that we must have reasons for our convictions: to believe something is to believe it to be true and if you don't have any grounds for thinking it true, you shouldn't believe it. For example, information installed in our brains as part of our genetic endowment may exercise a beneficial influence on our behaviour but such evolutionary 'memory' is not a repository of knowledge: to make it such, we must have grounds for relying on it. Nevertheless, we must also get away from the radical internalist idea that having a reason to believe is a matter of being able, at the present moment, to produce evidence.

One's belief may be well-grounded in past reasoning even if one is quite incapable of recapitulating that reasoning and there is no need to invent some alternative support for the belief which one can now bring to mind. We are not creatures of the moment, unable to carry our cognitive achievements forward from one instant to the next.

Justifying Memory?

Suppose I remember that Hitler committed suicide. I don't remember how I learnt this, nor can I lay my hands on anything that might count as (direct) evidence in favour of it. This is the situation we find ourselves in with the bulk of our factual beliefs: how do you know the boiling point of water or the dates of the First World War? It seems that either such knowledge is completely groundless or else it is based on the simple fact that you remember these things. But how can this memory be self-validating? How can a belief be evidence for its own truth?

As already noted, some philosophers have sought to argue that memory beliefs are self-validating. Either they think a memory belief is intrinsically self-validating, so that if one appears to remember that *p*, one already has a good reason to believe that *p*. Or else they appeal to background knowledge about how one is likely to have acquired the belief in question: I know that I learnt most of my history from a number of reliable sources – teachers, reputable books, etc. – therefore I know that I am likely to have got the belief that Hitler killed himself from a reliable source. Either way, the very fact that I have this belief can serve as evidence for its truth.

In this section, I shall discuss the latter view on which memory is a source of inductive evidence (without intrinsic authority) whose probative force rests on the grounds we have for thinking our memory to be a reliable guide to the truth.² Here we are supposed to start out with a neutral attitude to our own memory beliefs, to pretend that *a priori* we are as likely to be wrong as to be right in what we recall. But can we really regard our memory beliefs as we would the clicks of a Geiger counter? Our memory is not one more informational device which we can use or not as we please: it is fundamental to all cognitive transactions, including any that would be involved in establishing the reliability of memory itself.

Any investigation into the reliability of memory will make use of beliefs about the past.³ I may simply *recall* having learnt my history from trustworthy authorities. Or if I have little idea about the origins of my historical beliefs (or about the reliability of my sources), induction from past instances of memory success and failure may still assure me that my memory is to be trusted on historical matters. But I could hardly conduct such an investigation without using my memory to compare what I claimed to remember with what I later discover to be the truth. So in any such enquiry, some memories must be taken at face value, at least to

start with. An agnostic about memory could not even begin to determine which of his memories he should accept and which he should suspect.

My opponent may reply that I have ignored the possibility of an abductive investigation into the reliability of memory.⁴ Here, one takes all one's apparent memories and asks: what is the best explanation for the fact that I have all these memories? The answer may well be that I have the memories I do because the world was at least approximately the way I remember it as being. The details of such an argument need to be spelt out and until this is done, it is hard to judge whether a sound abductive inference could underwrite our memory knowledge. But however the inference goes one can be sure it will be quite complicated: we must consider a range of memories and a reasonable selection of possible explanations for them and assess each by our chosen standards of abductive inference before we can arrive at the conclusion that our preferred explanation is the best.

Memory will be involved at every stage here, not memory of the past successes and failures of memory but memory of what has already been established in reasoning: that such and such a sub-set of our apparent memories are well explained by a given hypothesis, that such and such is not as good an explanation of the whole set of apparent memories as the one we are now considering and so forth. I could hardly hold in mind all the stages in such a complex argument. If it is written down, then perhaps I can get my mind around any given stage of the argument at will; but my present sense that I could do this successfully if I tried must ultimately be based on my *memory* of what happens when I do try.⁵ So an abductive investigation of memory's reliability is as dependant on memory as an inductive investigation.

There undoubtedly are cases in which one learns things by utilising background knowledge about memory. Dennett asks how we would go about answering two questions: (1) Have you ever danced with a movie star? (2) Have you ever driven for more than seven miles behind a blue Chevrolet?⁶ Most of us, he says, would deliver a firm 'No' to the first question but would reply 'Don't know' to the second. Yet in both cases we can't remember anything directly relevant to the proposition. So why treat them differently? The reason, Dennett suggests, is that we think that had we danced with a movie star we would now remember the fact while we don't think this in the case of driving behind a blue Chevy. So our memory provides inductive evidence for the proposition that we have never danced with a movie star but not for the proposition that we have never driven behind a blue Chevy for some distance.

But is it always like this? Suppose we did remember having danced with a movie star. Must we have satisfied ourselves of the accuracy of our memories before accepting this proposition? Certainly, we might learn things which lead us not to accept the proposition: I might suspect that I was inclined to deceive myself on such matters. But in the absence of any evidence about how accurate my memory was in this regard, what would my epistemic position be? Would I feel entitled to accept the verdict of my memory that I had danced with a movie star or not? Surely yes. But then I am not neutral: my default reaction to memory is acceptance.

Memory as *Prima Facie* Evidence

In the face of such considerations, several philosophers have argued that memory's epistemic authority must be intrinsic. If I remember that Hitler committed suicide, then I am simply entitled to continue to hold that belief unless and until countervailing evidence is presented to me.⁷ To put it another way, my recollections have *prima facie* authority: I can trust my memory without the sort of supporting arguments sketched above. Something like this must be right, but it remains to elucidate '*prima facie* authority'.

On one reading of this phrase, what is being said is that memory gives us a form of evidence for *p*, '*prima facie* evidence', which establishes the truth of *p* in the absence of countervailing evidence. This would be to model the epistemic role of memory on perception. In perception, we have two separate states: the belief and the experience which furnishes *prima facie* evidence for the belief. But memory is an awkward fit for this model. When I remember that Hitler committed suicide, are there really two elements here: the belief that Hitler killed himself and the memory impression on which that belief is based? All I seem to find when I ask myself what I remember of Hitler's death is a series of beliefs. If my memory of Hitler's suicide were experiential (had I actually been present at the suicide and could visualise what happened) then we could have discerned two elements – the memory experience and the belief based upon it – but our topic is factual and not experiential memory. Factual memory is a mechanism for preserving beliefs already acquired: it is not a source of quasi-experiential evidence.⁸

Memory Without Belief?

Pollock wishes to restore the parallel between factual memory and sensory experience and with it the idea that memory provides evidence for belief. The belief-independence of sensory experience is established by considering cases in which we enjoy an experience as of *p* but don't trust our senses and therefore don't form the corresponding belief. In much the same way, Pollock thinks, we can have a memory without trusting it.⁹

Say I seem to remember that I was born on New Year's Day: that date is the one which comes to mind whenever I ask myself when I was born. Such a memory may persist long after I become convinced that my memory for such facts is unreliable: New Year's Day still pops into my mind whenever I am asked my date of birth. Here I have the memory but I don't trust it, so I don't believe what it tells me. My memory of my date of birth is not a belief but a state distinct from belief which, in Pollock's view, gives me *prima facie* evidence for a belief, evidence which is in this case undermined.

Pollock infers from such examples that even when I have full confidence in my memory, we must distinguish the state of remembering that *p* from the state of believing that *p*. Now I agree that there are phenomena, not beliefs, which may

be called memory impressions, but I have two objections to what Pollock claims about them. Firstly, it is phenomenologically very implausible to suppose that such memory impressions are present whenever we consult our memory: in (factual) memory there is nothing like the sensory impression which occurs whether or not we accept it. Secondly, such memory impressions as there are cannot play the fundamental epistemological role Pollock casts them in. Pollock's cases of memory without belief are best redescribed as cases in which we have a feeling, an impression or a hunch but no memory, a redescription which makes the (non-*prima facie*) evidential status of these impressions clear.¹⁰

Rushed for time by the quiz master, I am asked whether there are more than 120 members of the United Nations. My instant reaction is that there are and I plump for the answer 'yes', but I am far from claiming to know that there are more than 120 members. I have no doubt that I am under this impression because of various things I already believe about who is a member of the UN – my response does not strike me as a pure guess or an intuition of an *a priori truth* – but I don't think myself able to fix immediately on the right answer to such a precise numerical question.¹¹ In forming a belief on the matter I must forget about the hunch altogether and perform some sort of rough calculation, calling to mind what I do know about the membership of the UN.

Do I remember that there were more than 120 members in the UN here? My linguistic intuitions are not very clear on this point: perhaps such an impression can properly be called a memory, provided it is generated by beliefs acquired in the past. But what matters is not what name we give the impression but whether such an impression should be regarded as providing *prima facie*, non-inductive evidence that the impression is correct. I cannot see why it should. The only argument which might be offered on this score is that we have to attribute such a *prima facie* authority to memory-impressions if we are to avoid memory scepticism. But there is no need to grant this status to cognitive impressions which are not (current) beliefs in order to get on with our cognitive lives.

The true situation seems to me to be this. Whether my former belief that I was born on New Year's Day and the impression which it has left as a residue constitute evidence in favour of this proposition very much depends on my view of how reliable these beliefs and impressions are. They have no *prima facie* authority for me. It is hard to see how the fact that I once believed something should, in itself, provide me with any reason to believe it now. And it is equally hard to see how such a past belief could acquire a special evidential weight simply by persisting in the form of a current impression. I re-adopt past opinions, since abandoned, just in so far as I have grounds for thinking them an accurate guide to the truth.

Memory as Belief-Preserving

So Pollock's memory-impressions have no *prima facie* epistemic authority. But I cannot take such a non-committal attitude to my present beliefs. If I find myself with the belief that Hitler committed suicide, if I find myself convinced of this

point as a result of past cognitive activity, I can't regard this adherence simply as a more or less reliable indicator of what might be the case, or even as a piece of *prima facie* evidence. What memory preserves is belief itself and to believe that *p* is precisely to have finished enquiring into *p* by forming the view that *p*: it is not to be in possession of a sort of evidence for *p* (either *prima facie* or inductive) which, if not outweighed by contrary evidence, will convince one of *p*.¹²

Awareness of evidence motivates both the formation of belief and its abandonment, but it cannot motivate the maintenance of belief in memory. If I believe that *p* only because I regard my memory of *p* as evidence in *p*'s favour, then whether *p* must be an open question for me which I resolve in the light of my memory of *p*. But if at time *t* it is an open question for me whether *p*, then at time *t* I have no memory-belief in *p*. To remember that *p* is to no longer feel the need of evidence for *p* and so no memory can maintain belief by furnishing evidence for it.

True, I may note that my memory of Hitler's suicide provides me (or others) with evidence that Hitler killed himself. Here, I do not re-open the issue of how Hitler died but merely note a point in favour of a belief I would maintain anyhow. This may be in an effort to provide others with grounds for believing what I do. But the belief itself is seriously in question only if its retention depends on the success of this procedure and a belief is held only so long as it is not seriously in question. If doubt overwhelms me and forces the abandonment of this belief, I can then use the fact that I previously believed that Hitler committed suicide as inductive evidence in favour of his suicide. Such reflections on what I used to believe might even restore my conviction, but the restored conviction would be a different belief with a different justification.

Belief states, so construed, have an obvious utility. If rational belief in *p* required continued assessment of evidence for and against *p*, then not only would we have to keep an eye out for such evidence, we would also have to store the evidence we already have on the matter so that the significance of new evidence could be assessed in the light of it. But this is quite impractical given the number of issues we need to have a view about. Once a question is decided, we close the books on it and throw away the evidence: deliberately retaining evidence for future consultation is a sign of doubt, an attitude appropriate to the scientist who is interested in the likelihood of various things and has a professional obligation to suspend judgement but quite unsuited to the everyday believer.¹³

I have rejected the idea that memory provides *prima facie* evidence for belief, but I remain sympathetic to the claim that memory has some sort of *prima facie* authority. In the absence of specific grounds for doubt, we are entitled to persist in believing what memory serves up to us; not because memory provides evidence for these beliefs, but in another way.¹⁴ What way is that?

The Function of Memory

I have characterised memory as a faculty for preserving belief and belief as a state in which we feel no need of evidence for the proposition believed. But even if

those who remember do not feel impelled to ask why they should believe what they remember, we observers can still pose this question. We can ask how memory preserves the rationality of beliefs if not by providing the believer with continuing evidential support for his belief.

Say I am engaged in a rather long deductive argument. I believe p and q ; after some effort I prove that p and q together entail r and further exertions establish that r and another proposition I accept, s , together entail t . Therefore, I come to believe t . Now, as I am proving t , I make no effort to hold in mind the proof I discovered for r ; neither do I have time to review the proof of r once I have arrived at t . I may even have forgotten it. But since r has been established to my satisfaction, I feel entitled to use r in this and any future argument. What is going on here? I have already dismissed the view that I believe r because my subsequent mental state provides me with some kind of evidence that r is true. Rather, my grounds for believing r long after I have proved it are precisely those grounds which led me to believe r in the first place. Provided nothing has happened in the meantime which should make me doubt r , my continued belief in r is rational on just the same grounds as my original belief in it was.¹⁵

The core idea here is that memory is a faculty which preserves the probative and motivational force of evidence beyond the point at which that evidence has been forgotten. It enables current evidence to sustain and justify future belief by perpetuating its belief-fixing influence. When a belief is laid up in memory, the belief takes with it the probative force of the evidence which led us to adopt it in the first place. This is how memory enables us to retain knowledge previously acquired without our being in a position to rehearse the grounds on which we acquired it.

But how is the subject to think of his own epistemic situation once he has forgotten this evidence? What is it like for him when he claims to know via memory that p ? He believes that p and, like us, he thinks that to believe that p he must have some reason to think p true. Perhaps we, from the outside, can point to his earlier evidence as that which justifies his current belief, but is not the subject himself uncomfortably endistanced from the grounds for his own beliefs?

In fact, we can't suppose that our subject feels the need of evidence to support his continued belief that p , for to believe that p is precisely to have ceased to feel that need. All that will strike him from the first-personal standpoint is that he knows that p in a certain way, namely by remembering it. One seems to remember that p when one seems to have established that p at some point in the past and preserved that knowledge ever since. That is what makes a subject's memory beliefs feel different from those he has just acquired from testimony and other non-memory beliefs.¹⁶

Though memory does not (usually) act as a source of evidence for belief, the fact that one seems to remember that p is still relevant to whether one is justified in believing p . If one didn't appear to remember that p , if the belief in p seemed to have popped into one's head *de nouveau*, one wouldn't feel entitled to defer to

one's past self, to evidence one is no longer aware of, for a justification of that belief. To seem to remember that *p* is to feel entitled so to defer, provided one has no adequate reason for doubting one's memory.

Memory as Rationality-Preserving

A well-functioning memory preserves the rationality of belief, but not by preserving the evidence which prompted the acquisition of the belief. Rather, it does this by holding the belief in place with a force proportional to the strength of the evidence for it (a force which I shall call cognitive inertia). Given this, the subject will abandon the belief and resume enquiry into the matter just when he receives evidence sufficient to make doubt reasonable. Memory malfunctions where the tenacity of the preserved belief is out of line with, or is not rooted in, the earlier evidential support for it.¹⁷

There are many ways of fixing beliefs in place which do not preserve the rationality of those beliefs.¹⁸ Say that I am prone to entertain groundless doubts about the fidelity of my partner. Being momentarily free of such jealous suspicions, I decide, perfectly reasonably on the basis of the evidence before me, that my partner is not being unfaithful. But I know full well that I will abandon this conviction at the slightest prompting and fly into a jealous rage, so I decide to pay a visit to the hypnotist in an effort to fix this rational trust in place. Suppose the hypnosis works and my belief in my partner's fidelity is disturbed only by events which would arouse suspicion in any reasonable person. Does it follow that I have preserved the rationality of my belief, as well as the belief itself, by visiting the hypnotist?¹⁹

This depends on the answer to a further question: does the efficacy of the hypnosis depend on whatever evidential support I have for the belief? If so, the hypnosis is an aid to the rational retention of belief. But suppose the hypnosis works just as well regardless of whether I am inclined (in my lucid moments) to believe in my partner's fidelity: here the hypnosis does not preserve the belief's motivation, rather it replaces it. So even if I do have grounds for trust in my partner, and even if the strength of the hypnotically induced belief is proportional to the strength of the evidence I have for their fidelity, that evidence cannot justify my belief. The evidence motivates me to consult the hypnotist but it does not *directly* motivate the retention of the belief, so while it may help to justify the action required to fix this belief in place, it cannot justify the belief itself. A true aid to memory works off the probative force of the evidence which supports the belief, it does not replace it. Visiting the hypnotist may be a perfectly reasonable way of preserving a rational belief without thereby preserving the rationality of the belief.²⁰

We can now see the grain of truth in the claim that memory provides *prima facie* evidence for belief. Our memory has a *prima facie* epistemic authority in that we are entitled to persist in believing something remembered provided nothing comes to our notice which should make us desist. But this is because memory

preserves the rationality of that belief, not because it gives us (*prima facie*) evidence for the proposition believed.

Epistemic Conservatism

An epistemic conservative holds (in Harman's words) that 'a belief can acquire justification simply by being believed'.²¹ I considered one form of that position when I dismissed the idea that my belief in *p* might be a source of *prima facie* evidence for that very belief. But the conservative need not imagine that the mere fact of belief constitutes evidence for the truth of the proposition believed. What he may think is that my belief in *p* provides me with non-evidential grounds for continuing to believe that *p*. Can such a position be defended and how does it relate to my own view of memory?

According to one sort of epistemic conservative, belief is a cognitive commitment to the proposition believed. Consider practical commitments. A person makes promises and signs contracts for all sorts of reasons, but once the commitment is made, he has a reason to do what he has promised, a reason over and above any which might have led him to take on that commitment in the first place. Making a promise is not like visiting the hypnotist in an effort to ensure that one implements a difficult decision when the time comes; rather, it is a means of giving yourself an extra *reason* to carry out that decision. To do something simply to keep a promise is a perfectly rational procedure. Now suppose beliefs are epistemic commitments; wouldn't finding oneself committed to a certain proposition provide one with a reason to carry on believing it, a reason which is not evidence for the truth of the proposition believed?

Two points tell against the commitment model of belief. First, promises are made to other people and their social function is to render inter-personal interactions more predictable, but my beliefs are primarily for my own future consumption and only derivatively for other people's. The commitment model needs the idea that one can bind *oneself* with a promise, that one can give oneself an extra reason to \emptyset simply by promising oneself that one will \emptyset . Auto-promising is an unfamiliar procedure, but several philosophers have argued that taking a decision on future action is like making yourself a promise: it gives you a reason to carry out the decision, over and above the intrinsic desirability of the action decided upon.²² If so, perhaps intention-formation provides the correct model for belief.

But this view of practical decisions is surely mistaken, for if I decide that I will \emptyset and it then becomes obvious that \emptyset -ing would be a terrible idea, I don't have any reason to do it simply because I decided to do it.²³ And the analogous view about belief seems equally mistaken. One doesn't make a belief more rational simply by believing it. If a belief is irrational when adopted, it remains just as irrational while laid up in memory.

Secondly, I have accounted for memory's role as a knowledge-retention system by reference to the rationality-preserving character of memory. Unless the epistemic conservative denies that memory is rationality-preserving, these peculiar cognitive commitments (and the non-evidential reasons for belief which they generate) are simply not needed to enable memory to perform its epistemic function. It is quite enough to suppose that the past rationality of our beliefs will be reflected in their current normative status. By contrast, the conservative move of awarding each of our beliefs an extra point, as it were, simply for being believed seems both gratuitous and indiscriminate. Theoretical economy and normative intuition each tell against it.

Groundless Memory

In reply, the conservative might pose the following question: what if there were no adequate justification for the original acquisition of a memory belief? What if the believer finds themselves with a belief which they have no reason to doubt but which is, in fact, ill-grounded? We can't require that they know why the belief was acquired as that would simply defeat the point of memory. So are they entitled to carry on believing provided they have no intimation of this ungroundedness? Surely they often are, whether or not the belief constitutes knowledge. Yet if all memory does is to preserve the belief together with its original justification, how can they be entitled to believe it?²⁴

In the eyes of the epistemic conservative, this is why we need to suppose that the cognitive commitment implicit in belief has a normative weight which goes beyond the probative force of the evidence which led us to adopt the belief. Even when a vow is irrational, we have some reason to keep it; similarly, might I not be entitled to stick to my (unreasonable) decision that *p* is true until given good reason to abandon it? But, as we shall see, it is unnecessary to postulate epistemic commitments in order to explain why it might be reasonable for me to hold on to a belief that was formed irrationally.

I seem to remember that Hitler committed suicide at the end of World War Two, so when I read a report in a reliable newspaper saying that an elderly Hitler was spotted in a village in South America some years ago, I conclude that the witness must either be mistaken or else a liar. I don't seriously re-open the issue. Now suppose that at the time I formed this belief, there was a great deal of uncertainty about Hitler's fate; needing the security of the thought that this evil man was no more, I ignored the evidence that Hitler was still alive and in due course became convinced he was dead. Here, my belief was irrational. This belief is preserved in memory and much later it leads me to infer that the South American witness is misleading us. I have no recollection of how it was arrived at – to me now it is just like any other memory and has as much claim on me. Am I being irrational?

Two-Dimensional Rationality

There are really two questions here which it is important to keep apart. The first is 'what reason do I have to suppose that Hitler killed himself?' Very little, so my belief is unjustified. But we could also ask 'is it reasonable of me to reconsider my belief that Hitler committed suicide?' And the answer to this question might be 'No', however irrational the belief in question.²⁵ Note, the believer himself can't even raise the former question until the latter has been settled, for to seriously consider whether p is true in the light of evidence against p is already to have abandoned one's belief in p . Therefore, if it is reasonable for me not to reconsider my belief in Hitler's suicide, I can't be irrational in continuing to believe it, though the belief itself may be quite irrational.

One might wonder how anyone could rationally acquire or retain an irrational belief. Won't the fact that I am justified in accepting a belief by itself ensure that the belief accepted is a rational one? Take memory. Suppose I have a well-grounded confidence in what I appear to recall. Won't that suffice to ensure that my belief in Hitler's suicide is itself justified? I might be unable to produce good reasons for what I claim to remember, but if there are no grounds for doubting that I previously had such reasons, isn't this absence itself a piece of evidence which justifies the belief retained, as well as my retention of it? True, my original conviction might have been unreasonable, but if what is now sustaining that conviction is a well-grounded faith in the contents of my own memory, surely my current belief is justified, its sordid history notwithstanding?

Here, the objector moves from the premise that I am entitled to believe that Hitler committed suicide because I apparently remember it (in certain conditions) to the conclusion that my apparent memory of Hitler's suicide (in those conditions) must give me sufficient evidence for the truth of p . But this inference is unsound. The absence of grounds for doubt sufficient to undermine my right to rely on my memory of p cannot be converted into the form of inductive evidence which (when combined with the memory itself) justifies the belief that p . No such conversion can take place until we know what factors are likely to render memory unreliable – something we can learn only by using our memory. Only then can we be sure that there is no evidence for memory's unreliability. And, as I noted above, unless we are entitled to accept what memory tells us prior to *any* empirical investigation of memory's reliability, dependence on memory can never be justified.

Cognitive Inertia

Why might it be unreasonable of me to re-open the issue of Hitler's fate? Perhaps I don't have time to look into the matter now; it would be difficult to dredge up a fair sample of the relevant evidence, assimilate it, weigh it against the newspaper report and so forth. And even if I did have time to do all this, the prospects of forming a well-grounded view on the matter may remain slim. Thus there is

no point in reconsidering the belief. Needless to say, I will not arrive at this conclusion by deliberation – rather, the absence of doubt will be motivated by a tacit appreciation of such considerations. Still, why shouldn't the testimony of the South American lead me to suspend judgement on the point at least until I can investigate further? The answer is that beliefs have a certain degree of inertia, they resist revision.²⁶

The cognitive inertia of belief is a corollary of the rationality-preserving nature of memory. Where belief is rational, the inertial force of the belief is determined by the strength of the reasons which supported its adoption; where the belief is irrational, it is determined by some other factor. Either way, a belief, once acquired, constitutes a psychological obstacle to its own revision: if it didn't, it could never propagate the motivational force of the considerations which led to its formation in the absence of those considerations themselves. So this cognitive inertia, far from being a regrettable lapse, is essential to the rationality-preserving role of memory.

The claim that even rational beliefs have some inertia may seem very puzzling. We sometimes encounter resistance upon trying to rid ourselves of a belief which we have come to think of as unjustified, but isn't this always a symptom of irrationality? In so far as we are rational, shouldn't we find it easy to rid ourselves of any beliefs we ought to be having doubts about? But my point is precisely that one of the factors that we need to consider in deciding whether a subject ought to be having doubts is the strength of his belief. If the belief is strong and the grounds for doubt are relatively weak, then it is not rational for him to abandon that belief. And this is so even if the belief's strength is, unbeknownst to him, not well-grounded in past evidence. It can be unreasonable to reconsider an irrationally strong belief precisely because of the inertial force behind it.

If, on the other hand, the grounds for doubt are so strong that it is reasonable of our subject to wonder whether *p* is true, the fact that he used to believe that *p* should no longer be an obstacle to the resumption of enquiry; nor should it be a factor in his deliberations about whether *p* (unless it can be treated as a bit of evidence). The cognitive inertia of belief does not show up in the theoretical deliberations of a rational subject precisely because it operates to prevent such deliberation: once the deliberation begins, the obstacle has already been overcome.

My memory decays as time goes on, not just in that I remember less, but also because my memory becomes less reliable. Where memory works well, it registers the gradual decline in memory's reliability in an accompanying decline in the cognitive inertia of the retained convictions. The cognitive inertia of the beliefs of a rational person must reflect not only the strength of the original justification for their belief but also the likelihood that both the strength and content of a justified belief will be correctly preserved in memory. Preserving the rationality of a belief is not just a matter of accurately recording the strength of the original justification – features of the preservation mechanism must be taken into account as well.

The Transfer of Epistemic Responsibility

Internalists hold that beliefs are subject to reason and that believers are responsible for the rationality of their beliefs. I have argued that a believer is entitled to transfer responsibility for his current beliefs on to his earlier self. A rational belief requires evidence to justify it, but an awareness of that evidence need not be co-present with the belief it justifies. Provided I was once aware of this evidence, the responsibility of justifying my present belief has been discharged and the connection between rationality and responsibility on which the internalist insists has been preserved.

When one of my memory beliefs is unjustified, am I always to blame?²⁷ Obviously, I am to blame if I am not entitled to rely on memory at all in this instance. I am at fault in another way if I rightly lean on my earlier self when my past self can't justify the belief. In the Hitler example just described, my memory is in good working order and the transfer takes place, but my past self is unable to discharge its epistemic responsibilities and I am to blame for the irrational belief that I (rationally) preserve.

But what if my memory garbles the content of a perfectly reasonable belief and simply fails to link me up with a past self who could bear the epistemic responsibility for the new belief? For example, I commit the justified belief that my car registration ends with an 'A' to memory, but I end up believing that it begins with an 'A'. Am I still at fault because I have an unjustified belief? To say I am at fault seems unfair since my memory's malfunction is something I might remain quite unaware of. But if I am not at fault, we seem to have an unjustified belief for which no one is responsible, finally breaking the internalist connection between justification and responsibility.²⁸

Here I rationally preserve a belief whose lack of justification I have no reason to suspect. Learning the truth, I would acknowledge that I had no reason for this belief, that it was groundless. But am I to blame for this? No. This is a case where I, quite rightly, attempted to transfer responsibility for these beliefs on to my earlier self but the transfer fails: no one is responsible for these beliefs and so no one is to blame. Have we now abandoned internalism?

Failures of Reason

Even the most radical internalist must concede that justification is linked to responsibility only when there is someone who is capable of taking on that responsibility, someone who can be held to account for an unjustified belief. For example, it may seem to me that I am reasoning cogently when I run through a valid mathematical proof whilst under the influence of speed, but it is very unlikely that such mentation could justify a belief in the theorem proved – the proof is valid and it appears valid to me, but since my ability to discriminate good reasons from bad ones is much impaired by the drug, my conviction is unfounded.

Here I have an unjustified belief in the theorem; but unless I have grounds for suspecting my own sobriety, I am not to blame for this belief (perhaps someone slipped speed into my tea just a few moments ago without my noticing). It would be mad to insist that before I can accept any mathematical proof, I must first assure myself that my brain is working normally. Rather, this is something I am entitled to presume in the absence of significant evidence to the contrary (otherwise all mathematical reasoning would require empirical support).²⁹ So I am entitled to have an unjustified belief, provided I have no grounds for suspecting that I am incapable of carrying out the reasoning on which it is apparently based.

Returning to memory, a radical internalist might complain that (on my account) if memory fails without giving the subject any grounds for suspicion, he ends up with an unjustified belief, though he is in no way at fault. But this puts someone with a mangled memory in no worse a position than the subject convinced of the truth of a mathematical theorem while quite unaware of the drug he has just consumed. The latter has no grounds for doubting the cogency of his reasoning, but its unreliability undermines any justification he may think he has. What exposes the subject to this element of epistemic luck is not the fact, peculiar to memory, that he is relying on reasons of which he knows he is not aware. Rather, it is the more general point that a subject may be entitled to think that a belief can be justified when, in fact, it cannot.

But isn't there an important disanalogy between the two cases? Surely the drug impairs the rationality of the mathematical reasoner, but if I rely on a bad memory, I am not thereby irrational. In fact, the cases are parallel. Both involve a failure in the processes which enable reasoning to take place, but in neither case does this amount to irrationality. If I reason carelessly, or in a biased way, I remain capable of responding correctly to reasons of which I am aware: I just fail to do so. This is irrational and is blameworthy. What is not blameworthy is being *unable* to respond to reasons properly and that is the predicament of the drug impaired. It is also the predicament of those who are entitled to rely on a plausible but misleading recollection of the conclusions they arrived at in the past: a faulty memory cripples reasoning. The only difference is that the reasons the latter cannot respond to properly are reasons they know they are no longer aware of.

So what exactly do moderate internalists mean to rule out when they insist on a connection between justification and responsibility? Prominent externalists have maintained that the following situation can arise: a belief is unjustified, the person with the unjustified belief is perfectly capable of taking responsibility for it, he has not even attempted to transfer that responsibility to anybody else and yet he is not to blame for the belief's lack of justification.³⁰ For example, suppose I suffer from olfactory hallucinations but, having no reason to suspect this, I believe that there is a foul smell in the room when there is none. Many externalists would claim that since my sense of smell is unreliable, beliefs based upon it cannot be justified. Yet I have not tried to transfer responsibility for this belief to anybody else and I am perfectly capable of taking that responsibility on myself: my olfactory failings do not in any way impair my reasoning abilities. So how can I avoid the blame for believing that there is a foul smell?

At this point the externalist applies our distinction, arguing that the belief is unjustified but the believer, having no reason to suspect this, is justified in having the belief. But this is a step too far for an internalist like myself. The externalist has finally broken the connection between belief-justification and believer-justification and with it any link between belief-justification and responsibility. To restore that link we should insist, as internalists do, that this experience justifies my belief provided I have no reason to doubt it and regardless of contingent facts about the reliability of the mechanism that produced it.³¹ In this case, both my belief and I are justified.

Descartes and Locke, the founders of the internalist tradition in modern epistemology, both felt anxious about our habitual reliance on memory.³² Ideally, they thought, we shouldn't believe anything unless the evidence for it is before our minds, for only then can we take full responsibility for our convictions. But, reluctantly, they agreed that we must rely on reasons we can't now recall to justify our beliefs. I think this reluctance misplaced; the internalist can happily rely on past reasoning without evading his intellectual responsibilities.³³

David Owens
 Department of Philosophy
 University of Sheffield
 Sheffield S10 2TN
 UK
 d.owens@sheffield.ac.uk

NOTES

¹ Locke 1975, pp. 657–9; Harman 1986, Chapter 4.

² Foster 1985, pp. 106–16.

³ Among the many authors who make this point is Meinong 1973.

⁴ Peacocke 1986, p. 164; Harman 1973, Chapter 12.

⁵ Descartes 1985, p. 15, and Locke 1975, pp. 533–4 were both aware of this point. See also Pollock 1986, pp. 46–58 and Plantinga 1993, pp. 62–3.

⁶ Dennett 1991, pp. 146–7.

⁷ Pollock 1986, pp. 83–7 and Plantinga 1993 *ibid.* endorse this view. So do Harman 1986 *ibid.*; Burge 1993, pp. 157–65; Chisholm 1997, Chapter 2; Dummett 1992.

⁸ For a discussion of experiential memory, see Owens 1996.

⁹ Pollock 1974, pp. 188–96. (Pollock is criticising Malcolm 1963.)

¹⁰ I was helped to see this by Michael Martin.

¹¹ I might be hesitating because I am trying to decide whether I already believe that the UN has 120 members, i.e. whether I remember this. This would be a rather different sort of case.

¹² Moran 1997, p. 151.

¹³ Harman 1986, pp. 38–42 and 46–9. For a defence of this view of belief, see Owens 2000, Chapter 9.

¹⁴ Burge 1993, p. 481.

¹⁵ Burge 1993, pp. 462–5.

¹⁶ Pollock observes that there is a *phenomenological* difference between remembering that something happened to you in the past and believing that it happened on the basis of testimony. He infers that even factual memory must have an experiential element, an element which would still be present in the absence of belief. See Pollock 1974, p. 51. But it is clear from what I have just said that we can highlight the distinctive feature of memory beliefs without endorsing Pollock's conclusion.

¹⁷ Much more needs to be said about when exactly beliefs should be abandoned or retained. Consider a couple of questions raised by Josh Wood. Suppose I have a belief and adequate support for it to start off with, but then evidence mounts up against it which ought to lead me to query it. Irrationally, I don't query it, but then the countervailing evidence is shown to be fraudulent. Do I end up with a rational belief? I would think not. Obviously the evidence which is holding the belief in place is being given more than its proper weight if the belief is impervious to rational reconsideration. Alternatively, suppose that I start off with an inadequately supported belief but then evidence comes along which fully justifies it, evidence which I note with satisfaction. Does this make my belief rational? That depends on whether this new evidence becomes the basis for my belief or not, something we can discover by asking what would happen if this new evidence were to disappear. If I would then stick to my guns, then my belief is irrational all along since it is held regardless of the adequacy of the evidence for it.

¹⁸ It is a familiar point that not all strategies for acquiring true beliefs are rational methods of belief formation: Firth 1981, pp. 149–56. The same is true of belief retention.

¹⁹ For a practical parallel, see T. Pink's discussion of the difference between decisions and decision drugs in his 1996, pp. 93–100.

²⁰ Alternatively, perhaps I realise that I ought to form the belief that *p* but I can't bring myself to do it: neurotic doubts keep crowding in. A hypnosis which removed the neurotic basis for the doubts would be an aid to rational belief preservation, but a hypnosis which simply implanted the belief without working off the probative force of evidence available to the subject would not.

²¹ Harman 1986, p. 34.

²² Raz 1975, p. 66 and Velleman 1989. Adler 1996 applies this idea to beliefs. Van Fraassen 1984 endorses the commitment model of belief for rather different reasons.

²³ For further discussion of this point, see Bratman 1987, pp. 23–6 and Pink 1996, pp. 125–35.

²⁴ Harman 1986, pp. 33–41 poses this question.

²⁵ For a similar line of thought, see Foley 1993, pp. 109–12. Foley notes a parallel with practical decision making: see Bratman 1987, Chapter 5. As I noted in the last section, a failure to revise a belief in the light of new evidence which should compel such revision is a sign of its irrationality. My point here is a different one: such unwillingness may be a sign of the belief's irrationality without its being the case that the subject is irrational in retaining it.

²⁶ Similar points can be made about intentions. Even if my decision to holiday in South Africa rather than France was an irrational one, it may not now be rational for me to reconsider it.

²⁷ Here I ignore the case where the belief preserved in memory was derived from testimony. In Owens 2000, Chapter 11, it is argued that when relying on testimony I transfer the responsibility for justifying my beliefs to other people.

²⁸ I thank Michael Huemer for pressing this important objection.

²⁹ Burge 1993; Burge 1997, pp. 28–9.

³⁰ McDowell 1998, pp. 427–43 for example, suggests that we distinguish the question of whether a belief is a responsible one from the question of whether we have any reason for it and implies that victims of hallucination who have no reason to doubt their senses also have no reason to believe what they do, despite their evident doxastic probity. See also the distinction between strong and weak justification in Goldman 1988 and that between belief and believer justification in Bach 1985, pp. 251–3.

³¹ I suspect that Burge and I would part company at this point. Burge does seem to think that we are entitled to rely on perceptual experience only if perception is actually reliable (whether or not the subject has any reason to doubt its reliability). See Burge 1993, p. 478 and Burge 1997, p. 28. See also the distinction between justification and entitlement which he draws in Burge 1993, pp. 458–9.

³² Locke 1975, pp. 533–4 and 657–9, and Descartes 1985, pp. 15 and 218–20.

³³ Many thanks to Robert Hopkins, Mike Martin, Jennifer Saul, Peter Carruthers, Christopher Hookway and Michael Huemer for their comments and to audiences at Leeds, Glasgow, Edinburgh, St Andrews and Rutgers Universities.

REFERENCES

- Adler, J. (1996), 'An Overlooked Argument for Epistemic Conservatism', in *Analysis*, 56: 80–4.
- Bach, K. (1985), 'A Rationale for Reliabilism', in *The Monist*, 68: 246–63.
- Bratman, M. (1987), *Intentions, Plans and Practical Reasoning*. Cambridge, Mass.: Harvard University Press.
- Burge, T. (1993), 'Content Preservation', in *Philosophical Review*, 102: 457–88.
- Burge, T. (1997), 'Interlocution, Perception and Memory', in *Philosophical Studies*, 86: 21–47.
- Chisholm, R. (1977), *A Theory of Knowledge* (2nd edition). New Jersey: Prentice Hall.
- Christensen, D. and Kornblith, H. (1997), 'Testimony, Memory and the Limits of the A Priori', in *Philosophical Studies*, 86: 1–20.
- Descartes, R. (1985), *The Philosophical Writings of Descartes*, Volume I, trans. J. Cottingham, R. Soothoff and D. Murdoch. Cambridge: Cambridge University Press.
- Dennett, D. (1991), 'Two contrasts: folk craft versus folk science, and belief versus opinion', in J. Greenwood (ed.), *The Future of Folk Psychology*. Cambridge: Cambridge University Press.
- Dummett, M. (1992), 'Memory and Testimony', in his *The Seas of Language*. Oxford: Oxford University Press.
- Firth, R. (1981), 'Epistemic Merit, Intrinsic and Instrumental', in *Proceedings of the American Philosophical Association*, 55: 5–23.
- Foley, R. (1993), *Working Without a Net*. Oxford: Oxford University Press.
- Foster, J. (1985), *A. J. Ayer*. London: Routledge.
- Goldman, A. (1988), 'Strong and Weak Justification', in J. Tomberlin (ed.), *Philosophical Perspectives 2: Epistemology*. California: Ridgeview.
- Harman, G. (1973), *Thought*. Princeton: Princeton University Press.
- Harman, G. (1986), *Change in View*. Cambridge, Mass.: MIT Press.
- Locke, J. (1975), *Essay Concerning Human Understanding*, ed. P. Nidditch. Oxford: Oxford University Press.
- Malcolm, N. (1963), 'A Definition of Factual Memory', in his *Knowledge and Certainty*. New Jersey: Prentice Hall.

- McDowell, J. (1998), 'Knowledge by Hearsay', in his *Meaning, Knowledge and Reality*. Cambridge, Mass.: Harvard University Press.
- Meinong, A. (1973), 'Toward an Epistemological Assessment of Memory', in R. Chisholm and S. Schwartz (eds.), *Empirical Knowledge*. New Jersey: Prentice Hall.
- Moran, R. (1997), 'Self-Knowledge: Discovery, Resolution and Undoing', in *European Journal of Philosophy*, 5: 141–61.
- Owens, D. (1996), 'A Lockean Theory of Memory Experience', in *Philosophy and Phenomenological Research*, 54: 319–32.
- Owens, D. (2000), *Reason Without Freedom*. London: Routledge.
- Peacocke, C. (1986), *Thoughts: An Essay on Content*. Oxford: Basil Blackwell.
- Pink, T. (1996), *The Psychology of Freedom*. Cambridge: Cambridge University Press.
- Plantinga, A. (1993), *Warrant and Proper Function*. Oxford: Oxford University Press.
- Pollock, J. (1974), *Knowledge and Justification*. Princeton: Princeton University Press.
- Pollock, J. (1986), *Contemporary Theories of Knowledge*. London: Hutchinson.
- Raz, J. (1975), *Practical Reason and Norms*. London: Hutchinson.
- Van Fraassen, B. (1984), 'Belief and the Will', in *Journal of Philosophy*, 81: 235–56.
- Velleman, J. (1989), *Practical Reflection*. Princeton: Princeton University Press.