

Unreasonable knowledge

Maria Lasonen Aarnio

This is the pre-peer reviewed version of the following article: Aarnio, M. L. (2010), Unreasonable Knowledge. *Philosophical Perspectives*, 24: 1–21, which has been published in final form at <http://onlinelibrary.wiley.com/doi/10.1111/j.1520-8583.2010.00183.x/abstract>

Abstract:

It is common orthodoxy among internalists and externalists alike that knowledge is lost or defeated in situations involving misleading evidence of a suitable kind. But making sense of defeat has seemed to present a particular challenge for those who reject an internalist justification condition on knowledge. My main aim here is to argue that externalists ought to take seriously a view on which knowledge can be retained even in the face of strong seemingly defeating evidence. As an instructive example, I first discuss whether a theory on which knowledge is belief that is safe from error has the resources for accommodating defeat. I argue that beliefs retained in defeat cases need not be unsafe or true in some accidental way. I then discuss externalist strategies for explaining why we have incorrect intuitions about defeat. The notion of an epistemically reasonable subject plays a central role in my theory. Reasonable subjects adopt general strategies that are good for acquiring true belief and knowledge across a wide range of normal cases, but stubbornly retaining belief in the face of new evidence does not reflect such policies. I argue that though the methods employed by subjects who fail to adjust their beliefs in defeat cases may be perfectly good, they are not good methods to adopt, as their adoption is accompanied by bad dispositions. What emerges is a view on which a subject can know despite being unreasonable, and despite failing to manifest dispositions to know across normal cases. Unreasonable subjects are genuinely criticisable, but like almost anything, knowledge can sometimes be achieved in the absence of a good general strategy.

1. Defeat

It has been an important desideratum of epistemological theorising in the post-Gettier era that one's theory be able to accommodate the phenomenon of defeat, the losing of an epistemic status such as knowledge or justification in circumstances of a specific kind. In what follows I will be interested in knowledge defeat. The circumstances in which knowledge is putatively lost need not involve any losing of belief. And it is at least controversial whether they involve any losing of information or evidence. Here are a few paradigm cases:

Trick on Suzy

At a time t_1 Suzy comes to know that a certain object is red based on perception. There is nothing abnormal about her perceptual abilities or the lighting in the room. At a slightly later time t_2 a highly reliable and trustworthy authority tells her that the object is illuminated by peculiar red lighting, lighting that would make objects of any colour look red.¹

¹ Chisholm (1966: 48), Pollock & Cruz (1999: 44), and Pollock (1995: 41) give a similar example, though they are concerned with defeat of reasons, justification, or evidence, and not defeat of knowledge. Hence, they need not assume that the subject initially knows what colour the object is.

Fred's draws

Fred places exactly one red and one black ball into a bag. Based on perception, at a time t_1 he knows that there is a black ball in the bag. He then starts making draws with replacement, carefully observing the outcome of each draw. Throughout, he remains certain that the contents of the bag don't change. By a later time t_2 he has made ten thousand draws, each of which has produced a red ball.²

The widely shared verdict is that even if Suzy and Fred stick to their beliefs, in light of the new evidence these beliefs no longer constitute knowledge. In both cases the new evidence is misleading: Suzy acquires evidence for thinking that there is trick lighting in the room, even though there isn't, and Fred acquires evidence for thinking that the bag doesn't contain a black ball, though it in fact does. As a rough and ready starting point, I will take knowledge defeat to be the losing of knowledge as the result of acquiring misleading evidence of a suitable kind.³ By 'defeat cases' I will mean cases such as the above, intending to remain neutral on whether knowledge is in fact lost.⁴

Defeat appears to present a particular challenge for externalist theories of justification and knowledge.⁵ Why can't Suzy's and Fred's beliefs continue to be held by a perfectly reliable method? Indeed, some of the most powerful alleged counterexamples to externalism are cases of the sort most famously given by Bonjour that involve defeating evidence. Recall Samantha who sticks to her belief that the president is in New York City based on a completely reliable power of clairvoyance despite possessing masses of misleading evidence (news reports, allegedly live television footage, etc.) for thinking that the president is in Washington, D. C. Bonjour objects that in such circumstances, Samantha's belief is not justified. And he simply takes it for granted that in the absence of justification, neither can it count as knowledge⁶.

In response to such examples, externalists have done their best to accommodate defeat. Some have used brute force, building into their theories an extra

² From Williamson (2000: 205).

³ Some epistemologists are far more liberal about defeat, allowing for beliefs that have no positive epistemic credentials to count as defeaters (see, for instance, Bergmann (2006) and Goldman (1986: 62, 111)). One example of such a defeater would be a belief about trick lighting that is formed out of sheer paranoia. Unless any belief is part of a subject's evidence, such cases don't involve the acquisition of misleading evidence. I will set such cases aside because they are far more controversial.

⁴ It should be clear that defeat in the intended sense should be distinguished from what is often referred to as *propositional defeat*, the core notion in a research program dominant in the heyday of the Gettier debate. The aim of this research program was to disqualify Gettier-cases as cases of knowledge on the grounds that the subject's evidence doesn't satisfy a certain completeness condition. The original, crude idea was that knowledge of a proposition p is incompatible with the presence of true propositions such that, if they were known, or if they were part of the subject's evidence, they would defeat the subject's justification for believing p (See Pappas & Swain 1978: 28). Note that such a defeasibility analysis of knowledge is in tension with the idea that knowledge, and not just justification, is defeasible. For if having knowledge is incompatible with the existence of true propositions that would defeat one's justification, then knowledge cannot be defeated, at least not through justification defeat.

⁵ By "externalist theories of knowledge" I mean theories on which knowing doesn't require satisfying some sort of internalist justification condition.

⁶ Bonjour (1980: 59-60). Bonjour (1980) was arguing, in particular, against the sort of reliabilist theory of justification put forth by Armstrong (1973), but he seems to think that similar counterexamples could be constructed for any externalist condition on justification.

condition stating that the relevant belief is not undermined in any way.⁷ It is hard to see such strategies as more promising than ones trying to deal with Gettier-cases by appeal to a condition on knowledge stating that the subject is not in one of them. Others have tried more subtle strategies, framing their favoured externalist condition on knowledge in such a way as to make room for the idea that it ceases to be met in defeat cases. Though I find none of these strategies appealing, my purpose here is not to argue against them.⁸ Rather, I want to canvass a radical option not many externalists have taken seriously, which is denying the correctness of our intuitions about defeat. I sketch a picture for explaining why it is that we have defeat intuitions, and about why, though these intuitions are mistaken, subjects who retain knowledge in defeat cases are genuinely criticisable. My aim will be to paint a rather big picture, and I will be concerned more with a general strategy than the execution of its details.

The core thought to be defended is the following. Subjects who retain beliefs in defeat cases fail to act in an epistemically reasonable manner, but though reasonableness is connected with knowledge in an important way, it is not necessary for knowledge. Reasonableness is at least largely a matter of managing one's beliefs through the adoption of policies that are generally knowledge conducive, thereby manifesting dispositions to know and avoid false belief across a wide range of normal cases. Subjects who stubbornly stick to their beliefs in the face of new evidence manifest dispositions that are bad given the goal of knowledge or even of true belief. But we need evaluative notions that can be satisfied without possessing a general disposition to satisfy those notions. There is no reason to think that knowledge is special in this respect: sometimes being stubborn pays off.

Before sketching the broad outlines of my "indefeasibilist" epistemology, I will start by motivating what might seem like very drastic measures. In particular, one might wonder just why the externalist could not give a straightforward account of defeat after all. Perhaps beliefs retained in defeat cases always fail to satisfy some sort of externalist condition that pretty much everyone will have to accept, namely, a condition geared towards ruling out the problematic sort of luck present in Gettier-cases. Because I think a safety condition is the most promising attempt at formulating an anti-Gettier condition on knowledge, my focus will be on the question of whether beliefs retained in the face of misleading evidence are always (and necessarily) unsafe. Those unsympathetic to safety should view the discussion as a case study of the kinds of problems that straightforward externalist attempts to deal with defeat face, at least when executed within the context of simple theories of knowledge. What I hope will emerge is the following general conclusion: beliefs retained in defeat cases need not be true by luck, or true in some accidental way. At least those in pursuit of a simple, minimalist theory of knowledge should seriously reconsider whether defeat intuitions really are correct.

⁷ For instance, Goldman's (1986: 111-112) approach strikes me as being dangerously close to dealing with defeat by brute force. Goldman deals with defeat cases by building in a "nonundermining" clause to his theory of justification and hence, to his theory of knowledge.

⁸ What is perhaps the most noteworthy externalist attempt to deal with defeat in recent years is Michael Bergmann's (2006) defence of a proper functioning view of justification and knowledge that builds heavily on basic moves made by Alvin Plantinga. Bergmann argues, first, that if the cognitive faculties of a subject in a defeat case were functioning normally, she would come to believe that she has a defeater for the relevant belief, and second, that the belief cannot be justified in the presence of a belief that it is defeated. A strategy along these lines won't appeal to those dissatisfied with a framework making use of the idea of proper functioning.

2. An attempt at a straightforward externalist account of defeat (and why it fails)

I will start by considering a very simple externalist theory on which knowledge is belief that is safe from error.⁹ As a rough and ready approximation, a subject's belief is safe from error just in case she could not easily have held a relevantly similar false belief.¹⁰ For instance, if Suzy looks at a stopped clock that just happens to read the right time, her belief could easily have been false and hence, it is not safe from error. I doubt that a reductive account of relevant similarity can be given. But what is clear is that a significant determinant of relevant similarity is just how a belief was formed. It has become common to invoke *methods* in this connection, the idea being that for two beliefs or belief episodes to be relevantly similar, they must involve an application of the same method. It is worth saying a bit more about methods, since the notion of a method or rule of belief revision will play an important role in the discussion below. I think great caution is in order here.

Here is a tempting line of thought concerning how methods relate to relevant similarity of beliefs. I take it to be essential to the idea of a method that methods are types that token episodes or processes of belief formation fall under in virtue of exemplifying those methods. Then, the first part of the line of thought is that for each token episode, there is a unique salient method that that episode exemplifies. Methods are types that split belief-forming episodes into equivalence classes, which is to say that the relation of being an application of the same method is reflexive, symmetric, and transitive. The second assumption has to do with how methods relate to the issue of which beliefs or belief episodes are relevantly similar. The thought is that a method encodes the way in which a belief is formed, and being produced by the same method is a necessary condition for relevant similarity. These assumptions seem to be implicit in formulations of the safety condition on which a belief is safe just in case the method that produced it could not easily have produced a false belief. Similar assumptions also seem to be at work in formulations of the sensitivity condition that make use of the notion of a method.¹¹

The tempting line of thought incorporates two assumptions that cannot be reconciled: the assumption that methods split token belief-forming episodes into equivalence classes, and the assumption that any two relevantly similar belief episodes are applications of the same method. An immediate argument against combining these assumptions arises simply from the observation that when we talk about relevant similarity of beliefs, what we are interested in is a similarity relation. But similarity relations are not equivalence relations, since they are not transitive: *a* may be relevantly similar to *b*, and *b* may be relevantly similar to *c*, without *a* being relevantly similar to *c*. One might also question the symmetry of relevant similarity¹², but non-transitivity is enough to make the point. Of course, this is not to say that which method a subject employs in forming a belief is irrelevant for evaluating the epistemic status of that belief. I will assume that it is at least a necessary condition on safe belief and hence, on knowledge, that the method employed could not easily have

⁹ For such views, see Sosa (1999a, 1999b) and Williamson (2000).

¹⁰ Or, to deal with cases in which a subject could easily have undergone a belief episode that fails to express any proposition, one could instead formulate the condition as follows: a subject's belief is safe from error just in case she could not easily have undergone a relevantly similar belief episode failing to express a truth.

¹¹ See Nozick (1981).

¹² In particular, take an ordinary good case and a sceptical bad case. One might think that from the perspective of the bad case, the good case is relevantly similar, but not vice versa.

given rise to a false belief. With the above caveat about methods in place, let me turn to the safety theorist's strategy with defeat.

Take one of the paradigm examples of defeat described above, *Trick on Suzy*. The case involves first knowing that an object is red based on looking at it, but then being told by a reliable authority that it is illuminated by red trick lighting. For the purposes of the present discussion, I will take the details of the case to be as follows. Suzy first sees the object at a time t_1 , keeps looking at it throughout an interval of time including t_1 and a later time t_2 , and is told at t_2 by a reliable, trustworthy authority that the object has been illuminated by red trick lighting throughout the time she has been looking at it, lighting that would make an object of any colour look red. Suzy is certain throughout that the colour of the object does not change. When first seeing the object, she forms not just the belief that the object is red at t_1 , but the belief that the object is red throughout an interval of time that includes t_1 and t_2 . Let r be the proposition that the object is red throughout this interval. At t_1 Suzy comes to believe r based on perception. Upon receiving the misleading testimony, she acts dogmatically, continuing to believe r at t_2 . The result we want is that at t_2 her belief in r is no longer safe.

Let me first introduce a distinction between believing a proposition on the same basis as before and believing a proposition by employing the same method as before. Forming a belief at least often involves a token application of an epistemic rule or method, but one might think that such a process does not occur at any instant of time during which the belief is retained. I will express this idea by saying that by simply retaining a belief a subject continues to believe p on the same basis as before.¹³ This should not be confused with believing p based on the same method as before, since a subject might undergo a new belief-forming process that exemplifies the same method as the one she previously used. Assume, for instance, that Suzy comes to believe that Tuesday is a rainy day based on looking outside in the morning. She then immerses herself in a book, ceasing to consciously entertain the belief, and forgetting all about the rain. At noon she takes a look outside again, and undergoes a new mental process producing a belief that Tuesday is a rainy day. In this case, she may well believe that Tuesday is a rainy day by using the same method as before, but not on the same basis as before.

Returning to *Trick on Suzy*, because Suzy originally knew r , at t_1 it must have been the case that her belief in r was safe. Hence, at t_1 she could not easily have falsely believed the object to be red. Now, if Suzy simply retains a belief in r on the same basis as before, it is difficult to see how her later belief could fail to be safe. For presumably, it remains true throughout that Suzy could not easily have initially formed a false belief about the colour of the object. Moreover, there is no reason to think that the mere retention of her belief introduces a failure of safety, assuming that Suzy could not easily have, at t_2 , ended up believing the relevant object to be any colour other than red. Whether she could easily have given up belief in r seems to be irrelevant, in so far as safety requires merely being safe from error, not being safe from failing to hold a true belief.¹⁴ The lesson is this: if Suzy continues believing r on

¹³ Even if retaining a belief changes its basis, we can think of the new basis as at least including the original basis. For instance, instead of believing a proposition just based on perception, perhaps one believes it based on retaining a belief originally based on perception in memory.

¹⁴ There are other defeat cases in which the safety theorist might attempt to explain the losing of knowledge by the fact that the subject could easily have switched from believing a proposition p to believing a proposition $\sim p$. For instance, take a case in which a subject initially knows p , but then acquires very strong evidence for believing $\sim p$, evidence that would make a belief in $\sim p$ reasonable.

the very same basis as before, then as long as the mechanism by which she retains her belief is such that it could not easily have led her to believe the object to be any colour other than red, her belief in r is safe at t_2 as long as her belief in r is safe at the earlier time t_1 – which by assumption it is.

In light of this, here is what the safety theorist's general strategy with defeat cases will have to look like. First, when putatively defeating evidence is acquired, the subject can no longer hold a belief on the same basis as before. She must rebase her belief by undergoing a token process exemplifying some method of belief-formation. Second, once the new evidence has been acquired, there is no method available to the subject by which she could come to form a safe belief in p . A more general idea emerges: the token processes giving rise to beliefs in defeat cases are always flawed in a way that prevents the resulting beliefs from counting as knowledge.

Let me first say a bit about how I am thinking about the availability of a method or epistemic rule. If a method telling one to believe p in circumstances C is not available to a subject, then for some reason the subject cannot successfully follow that method: she cannot come to believe p as a response to being in circumstances C .¹⁵ This might be the case, first, simply because she is not in the right circumstances and hence, cannot do anything as a response to being in them. But, second, even when in the right sorts of circumstances, she might not have the ability or right know-how to employ the method. Employing a method is following a rule, and the ability to follow a rule is not trivial. In the end, it may be that the repertoire of methods that ordinary humans are able to employ is surprisingly limited.

For the above idea to work, it looks like the method originally employed by Suzy can no longer be available for believing r once the defeating evidence is acquired. For if Suzy's original perceptual method is still available, and it produced knowledge of r at the earlier time t_1 , why would it not still do so? The safety theorist cannot exploit the thought that if a reliable person says that there is trick lighting, then there could easily have been trick lighting, for if this was what prevented the belief formed at t_2 from counting as knowledge, it would also prevent the belief formed at t_1 from counting as knowledge. I argued above that sometimes cases in which a subject doesn't employ a given method will have to be counted as close to cases in which she does. But it's difficult to see how this could help, for there need not be any other candidate method that Suzy could easily have employed, a method that might have led her to a relevantly similar false belief. Another thought is that now the fact that misleading evidence is present enters into determining relevant similarity of beliefs or belief episodes: the question to ask is whether Suzy could easily have formed a false belief based on perception *in the presence of relevant misleading evidence*. But this seems to be conceding that the presence of misleading evidence is relevant for how the subject's belief was formed. It looks like a variant of a view on which the original method is not available, since now the presence of misleading evidence is characteristic of the method employed. I now turn to this idea.

Even if the subject dogmatically sticks to her belief in p , perhaps she could very easily have come to believe $\sim p$ instead, thereby believing a falsehood.

¹⁵ There is an issue about whether a subject could be said to truly follow a rule, method, or policy, even if she fails to conform to it. Boghossian (2008), for instance, states that conforming to a rule is not a necessary condition on following it. To avoid confusion, I speak of *successfully* following a rule or method, where successfully following a rule entails conforming to the rule. However, it certainly strikes me as odd to say that a subject could form a belief by employing a method telling her to believe p in circumstances C even if she doesn't in fact believe p as a response to being in circumstances C .

Why might the original method that produced a safe belief in r no longer be available to Suzy? Let me focus on the second source of the unavailability of a method, namely, the lack of right sorts of circumstances for its applicability. Call methods that are not applicable in the presence of defeaters *defeat friendly*. One example of a defeat friendly method is inference to the best explanation: the presence of misleading evidence will typically mean that p no longer best explains the data. For instance, that an object is red doesn't seem like the best explanation of data consisting of a reliable person saying that the object is illuminated by trick lighting and of the object looking red. The thought now is that ordinary perceptual methods are not applicable in defeat cases, and that they should be characterised in a way that incorporates a condition stating that no defeating evidence is present. Indeed, it is not at all implausible that the perceptual methods we normally employ are sensitive to whether defeating evidence is present: most of us are only disposed to believe things to be how they perceptually seem to be in circumstances not involving, for instance, evidence for thinking that one's senses are not to be trusted.

There is little doubt that we employ defeat friendly methods. And in any world broadly like we take the actual world to be employing methods encouraging one to adjust beliefs in the light of countervailing or undermining evidence makes excellent sense. But unfortunately, appeal to defeat friendly methods cannot provide a general account of why subjects in defeat cases lose knowledge, for it is certainly possible to retain belief in the face of defeating evidence. This is precisely what actual subjects sometimes do. Even if it had been established that Suzy's original method is no longer available to her in the new circumstances involving defeating evidence, what would be needed is the additional conclusion that *no* safe method is available. So let me finally look at the kinds of methods that subjects who dogmatically stick to their beliefs in the presence of defeating evidence might employ. In particular, I want to see how those methods could be made to look bad.

Here is an idea already alluded to above. When subjects retain beliefs in defeat cases, the presence of defeating evidence enters into a characterisation of the method employed. For example, perhaps the method employed by Suzy at the later time is believing p based on its perceptually seeming to one as if p in the presence of good evidence for thinking that in the present circumstances its perceptually seeming to one as if p is not a good guide to whether p . Now, one must concede that such a method is not *good* in the following sense: it does not reliably produce true belief or knowledge across a range of normal cases, since only too often successfully following it will lead to believing falsehoods. After all, a lot of the time evidence of a defeating kind is not misleading, and when one's senses are not to be trusted, relying on how things perceptually seem is likely to result in false belief.

However, its not clear how this observation helps the safety theorist, as following a method that is good in the above sense doesn't appear to be necessary for meeting the safety condition: an application of a bad rule doesn't always lead one into trouble. In particular, circumstances can conspire so as to make applications of generally unsafe or unreliable methods yield safe beliefs. For instance, if reading a certain barometer is to be a good or reliable method, the barometer has to give the right readings in a wide range of normal circumstances, including different air pressures. But safely producing true belief only requires being safe on a particular occasion of use, delivering the right results given inputs that could easily have ensued, inputs that are relevantly similar to the actual ones. A barometer might, for instance, have a fault that causes it to give inaccurate readings in anything but low air pressures. But if I operate solely in high altitudes, taking readings in situations in

which the air pressure simply could not easily have been anything but low enough for the barometer to give an accurate reading, beliefs formed by reading the barometer can be perfectly safe. More generally, beliefs formed by employing defective methods need not be true in a lucky or accidental way.

To sum up, I have found no compelling reason to think that beliefs retained in defeat cases are always unsafe. One reaction to this would be to concede that safe belief is not sufficient for knowledge: the method employed by a subject shouldn't just be evaluated from the perspective of whether it produces safe belief on that occasion, but from the perspective of whether it does so more generally. The thought would be that to produce knowledge, a method must also be good in the sense described above, yielding true or safe belief reliably, across a wide range of normal cases.¹⁶ And so the new theory of knowledge would state that knowledge is belief that is safe and produced by a good method, a method that generally tends to produce safe belief. Or, even better, knowledge is belief that is safe because it is produced by a good method.

Below I discuss a general consideration against such views of knowledge. For now let me just say that I wouldn't want the viability of an externalist strategy with defeat to rest on the claim that subjects in defeat cases always follow bad methods. In particular, I want to leave open the possibility of employing methods that are *success-entailing*, methods that produce true belief and perhaps even knowledge whenever successfully followed. An example of such a method would be the method of believing p when one sees that p . Or, to take an example that does not rest on an externalist individuation of methods, consider the method of believing that one is cold whenever one is in fact cold. Someone might object that actual subjects only ever form beliefs in response to how things seem to them to be, but then consider the method of believing that one seems to be cold whenever one in fact does seem to be cold. If subjects can follow success-entailing methods in the first place, I see no reason why such methods could not be followed in defeat cases. The method of believing p when one sees that p in the presence of evidence for thinking that one's senses are not to be trusted is just as success-entailing as the method of believing p when one sees that p . Or, consider the method of believing that one seems to be cold when one seems to be cold in the presence of evidence for thinking that one is not to be trusted about how things seem to one to be. Such methods or rules are perfectly good in the sense characterised above.

One of my aims in the rest of the paper is to argue that our intuitions about defeat don't track how good or bad the method employed by a subject is, but the dispositions and abilities manifested by a subject who follows that method. A subject can manifest bad dispositions by following perfectly good methods. Faulting subjects who follow good methods requires moving beyond a picture on which a subject's belief is to be evaluated in terms of whether the method she employs passes tests such as being generally safe or reliable. In particular, some perfectly good rules are such that adopting and following them is accompanied by dispositions that are bad in leading to false belief on other not at all abnormal occasions. Take, for instance, the rule telling one to believe p when one sees that p in the presence of evidence for thinking that one's senses are not to be trusted. The problematic disposition accompanying the adoption of this rule is a disposition to believe p when it's merely as if one sees that p in the presence of evidence for thinking that one's senses are not

¹⁶ To avoid circularity, the proponent of such a view could not characterise the goodness of a method in terms of its conduciveness to knowledge, but rather, would do better to focus on true or safe belief.

to be trusted. This disposition leads to trouble on other occasions in which it might be manifested, occasions that are not at all abnormal.

Having set aside a straightforward externalist theory of knowledge defeat, I now turn to strategies for explaining away defeat intuitions. After looking at a couple of alternative strategies, I will sketch my preferred view.

3. Explaining away intuitions about defeat

The sort of view that I want to defend denies that lacking evidence of the putatively defeating kind is a necessary condition on knowledge. Of course, this does not entail that whenever subjects manage to retain belief in defeat cases – something that in itself requires a non-trivial ability – they thereby retain knowledge. Some defeat cases will involve the violation of one's favoured anti-luck condition on knowledge. For instance, if Fred could easily have gone from believing that there is a black ball in the bag to believing that there are no black balls in the bag, this in itself might disqualify him from retaining knowledge, since in the new situation involving defeating evidence he could easily have held a relevantly similar false belief. But above I expressed severe doubts about whether similar manoeuvres are always available.

On the view I want to defend, intuitions about knowledge being lost in defeat cases are at least sometimes incorrect. In this respect, the view is error-theoretic. However, this aspect of the view is far less problematic than might at first sight appear, for, as I argue below, one can give a systematic explanation of defeat intuitions from within an externalist framework. Moreover, what explains our error in defeat cases is closely connected to the basis on which we make correct judgments in normal, paradigm cases of knowledge. There is a feature of ordinary cases of knowledge that is absent in defeat cases, and in ordinary cases that very feature typically enters into an explanation of why and how subjects manage to have knowledge. Hence, the source of our error in defeat cases is closely connected with what enables us to make correct judgments in ordinary cases.

Let me finally turn to how defeat intuitions could be explained from within an externalist framework. To meet the explanatory challenge, the externalist will have to do two things. First, she will have to find some feature recognised by an externalist theory of knowledge that is present in defeat cases but absent in more ordinary cases of knowledge, and second, she will have to say why we are reluctant to attribute knowledge in the presence of this feature. Before turning to my preferred view, I will briefly discuss two alternative ways of implementing this general strategy. According to the first, the required feature is ceasing to know that one knows. According to the second, the feature is its becoming improbable on one's evidence that one knows.

On the first suggestion defeating evidence always leads to some losing of knowledge, namely, losing the knowledge that one knows the relevant proposition p . Or, more generally, the claim would be that at least some iteration of knowledge is always lost. Note that this strategy only works as long as there are no defeat cases – cases in which we have the intuition that knowledge is lost – in which a subject starts out with only a single iteration of knowledge. Moreover, if knowledge of a proposition p can be retained even in the face of strong evidence that has (putative) defeating force with respect to one's belief in p , why couldn't any number of iterations of knowledge be retained? The suggestion rests on the viability of the claim that certain propositions simply cannot be known in the presence of a certain sort of evidence, but this is precisely what I have expressed doubts about above. The general

sorts of problems that arise, for instance, for the claim that defeating evidence always has the effect of rendering a belief in the relevant proposition p unsafe can also be expected to arise for the claim that such evidence always has the effect of rendering a belief in the proposition that one knows p unsafe.

On the second suggestion the effect of defeating evidence is not that a subject no longer knows that she knows, but that it becomes unlikely on her evidence that she knows. And – the thought goes – we tend to confuse its being unlikely that one knows with not in fact knowing.¹⁷ As far as certain paradigm cases of defeat are concerned, this suggestion looks promising. Take, for instance, *Trick on Suzy*. Assume for now that though Suzy starts out knowing that the object she is looking at is red (proposition r), she doesn't have further iterations of knowledge. And let us grant that before acquiring the defeating evidence, the probability on Suzy's evidence that she knows r is not low. This doesn't seem implausible. Presumably, Suzy is aware of the fact that her belief in r is formed by perception. She has no reason to think that anything abnormal is going on, and she knows that perception has a good track record of producing knowledge of the colours of ordinary objects in ordinary circumstances. But upon receiving the relevant testimony, Suzy acquires evidence making it likely that there is red trick lighting in the room. The relevant subset of her new evidence will consist of something like the following:

[r is true.]¹⁸

I believe r on the basis of perception.

A highly reliable authority tells me that there is trick lighting in the room that would make it perceptually seem to me as if r was true even if it wasn't.

It is not at all implausible that on such evidence, the probability that Suzy knows r is low, even if r itself is part of Suzy's evidence. After all, it is highly likely that there is in fact trick lighting in the room. And arguably, the probability of knowing r conditional on believing r on the basis of perception in the presence of trick lighting is low. For if a subject forms a perceptual belief in r in the presence of trick lighting, then her belief is not responsive to the colour of the object in the right sort of way. Had the trick lighting been any other colour, her belief would have been false.

But there are at least two problems with the proposed strategy. The first has to do with cases in which a subject starts out with numerous iterations of knowledge. In so far as the first strategy discussed fails, it is at least possible for the subject to retain all of these iterations. To deal with such cases, the second strategy can be generalised to state that in defeat cases at least *some* iteration of knowledge becomes improbable. But now the proposed explanation of intuitions about the losing of first-order knowledge starts to look dubious. Why might we have a tendency to confuse the improbability of, say, having fifth-order knowledge with lacking first-order knowledge? The second worry for the proposed strategy has to do with other paradigm cases defeat, in particular, ones commonly categorised as cases of overriding defeat.

Consider *Fred's draws*. Fred starts out knowing that there is a black ball in the bag based on perception, but then a probabilistically freakish sequence of events takes place, and he draws out the red ball 10 000 times in a row. Let p be the proposition

¹⁷ Williamson (2009) uses a similar strategy to explain intuitions about failures of multi-premise closure.

¹⁸ r will be part of Suzy's evidence if one thinks that everything a subject knows is part of her evidence.

that there is a black ball in the bag (and that there has been one there throughout the relevant interval of time). I want to leave open a view on which everything known by a subject is part of his evidence. Then, assuming that knowledge is not lost, the relevant portion of Fred's evidence will consist of something like the following:

[p is true.]

I came to believe p based on perception, by looking into the bag before the sequence of draws, and still believe p .

The sequence of draws made produced the red ball 10 000 times in a row.

The contents of the bag have remained the same throughout.

Why is it unlikely on this evidence that Fred knows p ? In *Trick on Suzy*, the new evidence made it likely that if true, the truth of Suzy's belief was accidental or lucky. But nothing similar appears to be going on here. The fact that the bizarre sequence of draws was made doesn't in any way undermine the original basis of Fred's belief. And on the sort of simple externalist view of knowledge I am defending, it's not clear why, for instance, merely the presence of defeating evidence makes it likely that knowledge isn't had. Now, assuming that p is not part of Fred's evidence, it does seem unlikely on Fred's evidence that he knows p . But note that this would be the case simply because of the improbability of p itself: given the sequence of draws, it is extremely unlikely that there is a black ball in the bag. In so far as p is unlikely – or we take it to be unlikely – on Fred's evidence, what strikes me as a more plausible hypothesis is that our intuitions are a response to p being unlikely on Fred's evidence.

Though appeal to the idea that in defeat cases it becomes improbable on a subject's evidence that she has knowledge looked initially promising, I have expressed doubts about explanations of defeat intuitions that build on this idea. I now turn to my preferred strategy, one appealing to the notion of an epistemically reasonable subject. As indicated in the beginning, my interest is in the broad outlines of the strategy rather than in its precise implementation. My explanation of intuitions in defeat cases will make use of the notion of an epistemically reasonable subject, and the related notion of a reasonable belief. I take such notions to play a central role in our practices of epistemic evaluation. For instance, subjects who form beliefs by sheer guesswork don't act reasonably, and neither do subjects who fail to form beliefs about their surroundings when in suitable perceptual states and in optimal perceptual circumstances. And what is of particular importance is that subjects who retain beliefs in defeat cases don't act reasonably. I argue below that reasonableness can be characterised in a way that is thoroughly externalist, and that defeat intuitions can be explained by giving an account of why we tend to confuse failing to act as a reasonable subject would act, or failing to have a reasonable belief, with failing to have knowledge.

4. Reasonable subjects

On the view I want to defend, the feature distinguishing defeat cases from more ordinary cases of knowledge is that subjects who retain belief in defeat cases act in an epistemically unreasonable manner. Reasonable subjects are not subjects who just happen to go about acquiring a lot of knowledge. Rather, they are subjects who manage their beliefs in a way that makes sense given the goal of knowledge acquisition. I will think of a subject's way of managing her beliefs as encoded by a policy of belief revision consisting of a set of sufficiently general epistemic rules or

methods. Then, the thought is that the policies adopted by reasonable subjects are ones that are *knowledge conducive* for those subjects. As will become clear below, just how knowledge conducive a policy is for a subject depends on the dispositions that accompany the adoption of that policy. Hence, the notion of reasonableness is closely connected with manifesting dispositions of a certain kind: reasonable subjects manifest dispositions that are good to have for the purpose of acquiring knowledge. That is, they manifest general dispositions to know.¹⁹

Reasonable belief can be characterised in terms of the notion of a reasonable subject: a belief is reasonable just in case the subject who holds that belief is being reasonable, that is, doing what a reasonable subject would or might do. Reasonable beliefs are formed and retained through the manifestation of knowledge conducive dispositions and abilities. But it should be clear that reasonable belief is not a necessary condition for knowledge. Indeed, one of the main claims I want to defend is that a subject can know without being generally disposed to know. Not only can the best of strategies sometimes fail, but most goals can sometimes be achieved in the absence of good strategies. Knowledge is no different in this respect. However, as I argue below, circumstances in which one knows whilst failing to manifest knowledge conducive dispositions are deviant in an important way. This will be important for why we tend to judge that a subject lacks knowledge in cases in which we take her to be epistemically unreasonable. But let me first say more about the idea of knowledge conduciveness and in particular, about what it is for a policy to be knowledge conducive for a subject.

First, when evaluating whether a policy of belief revision is knowledge conducive, one must weight the value of knowledge against the disvalue of, in particular, false belief. Dispositions that encourage profligacy in belief-formation, resulting in masses of knowledge but even greater masses of false beliefs, are not knowledge conducive in the intended sense. One might instead work with some idea of a knowledge ratio, but even this suggestion is too crude, for it doesn't take into account the fact that certain dispositions and abilities are highly cautious, only resulting in belief on rare occasions. It is at least doubtful whether such caution is good from the perspective of the goal of knowledge. I will propose no algorithm for how to weigh all of these factors, but talk of knowledge conduciveness should be understood in a way that allows taking them into account.

Second, just how knowledge conducive a subject's way of managing her beliefs is will depend on the possible world one considers. For instance, believing things to be the way they perceptually seem to be will fare far better in normal, non-sceptical worlds than in worlds involving systematic deceit. Reasonableness could be tied to the features of the world a subject in fact happens to be in. Though the basic

¹⁹ My favoured characterisation of epistemic reasonableness is in terms of knowledge rather than truth conduciveness. Whether one chooses knowledge or truth will often not make a difference: truth conducive ways of managing one's beliefs also tend to be knowledge conducive, and vice versa. However, the difference between the two perspectives is easiest to see when considering the question of whether it is reasonable to believe that one's lottery ticket will lose. In so far as it is not possible for a subject to know, merely based on the odds, that her lottery ticket will lose, a rule telling her to believe that her ticket will lose based on the odds is not knowledge conducive. On the other hand, the rule in question would seem to be truth conducive, as it would result in masses of true beliefs and only very few false ones. I find the conclusion that it is not reasonable to believe one's lottery ticket to lose merely based on the odds to be a plausible one, but the type of explanation of intuitions about defeat that I sketch below is compatible with thinking about reasonableness in terms of truth conduciveness. This is because the policies employed by subjects who stick to their beliefs in defeat cases are neither knowledge nor truth conducive.

gist of my explanation of defeat intuitions won't rest on this, my preferred view is instead to characterise reasonableness in terms of policies that are knowledge conducive in normal worlds, worlds broadly like we take the actual world to be in certain epistemically relevant respects.²⁰ Normal worlds don't involve masses of envatted subjects. Sources of belief that we take to be trustworthy and reliable (perception, testimony, memory, etc.) are trustworthy and reliable in normal worlds. However, not every contingency we take to hold in the actual world is shared by normal worlds. For instance, even if Suzy is generally known to be far more reliable than other subjects, this isn't true of any normal world. Trusting Suzy no matter what other subjects testify to without any independent reason to think Suzy to be more trustworthy is not reasonable. I take this parochialism to fit well with our intuitions about reasonable belief. For instance, brains in vats are reasonable in forming beliefs about their surroundings based on how things seem to them. Note that a similar result could be achieved by a view appealing not to policies that are in fact knowledge conducive (in normal worlds), but to ones that a subject reasonably believes to be knowledge conducive.²¹ However, suffice it to say that as an attempt to spell out the idea of reasonableness, such an account is viciously circular: the thought was that beliefs managed by following certain kinds of rules or methods are reasonable, but the right kinds of methods are now being characterised by appeal to reasonableness.²²

The account of reasonableness sketched incorporates a rule consequentialist way of thinking. Reasonableness is at least largely a matter of adopting or being guided by policies, thought of as sets of rules, that have certain consequences for knowledge. However, it is important to underline that I am evaluating rules and policies from the perspective of the consequences that result from their adoption by the relevant subjects. Such evaluations are importantly different from asking whether a rule or policy is *good* in the sense discussed above. Good rules rarely produce false belief when successfully followed. But as I shall now argue, a subject following a perfectly good rule might nevertheless be unreasonable, for adopting such a rule as part of her policy might result in an overall strategy that fails to be knowledge conducive. In other words, not any good method is one that a reasonable subject would adopt or follow.

5. Good methods and good methods to adopt

I want to contrast the question of whether a method is *good* with the question of whether a method is a *good method to adopt* for a particular subject. A discussion of exactly what it is to adopt or follow a method or rule is beyond the scope of this essay. However, let me say that as I am thinking about it, adopting or following a rule is closely connected with manifesting certain dispositions. In particular, if a subject adopts a rule telling her to ϕ in circumstances C, then other things being equal, she is

²⁰ Here one will be reminded of Goldman's (1986) theory of justification that likewise makes use of the notion of a normal world.

²¹ For instance, Wedgwood (2002) argues that a subject ought to (and that it is rational for her to) follow rules that she is rational in regarding to be "sufficiently reliable in the circumstances".

²² One of the reasons I have to be sceptical of the account is that it threatens to have über-internalistic consequences. Presumably, in some circumstances it can be reasonable for a subject to believe, for instance, that an inference rule such as affirming the consequent is valid (imagine a situation in which all the top logicians testify to the validity of the rule, and the subject is manipulated so that it seems valid to her). However, the conclusion that beliefs formed by employing the rule in such circumstances are reasonable strikes me as highly implausible.

disposed to φ as a response to being in circumstances C. But what will be important is that a subject with this disposition will also have other dispositions, in particular, the disposition to φ when she merely seems to be in circumstances C. The thought will be that when evaluating whether a method is a good method for a subject to adopt, one must look at the overall set of dispositions that accompany the adoption of that method.²³

So, for instance, a subject who adopts a rule telling her to believe p when she sees that p is disposed to believe p in circumstances in which she sees that p .²⁴ But at least given the cognitive makeup of normal subjects, having this disposition also entails having the disposition to believe p when one merely seems to see that p . More generally, having the disposition to believe p when one sees that p entails having the disposition to believe p whenever one seems to see that p , whether or not one is in fact seeing that p .²⁵ Having such a disposition means being susceptible to forming false beliefs in cases in which one's senses are deceptive. But this doesn't make the method of believing p when one sees that p a bad method to adopt, for such susceptibility to deceit is the price that must be paid for knowledge based on perception, and perhaps for any knowledge of the external world. After all, a disposition to believe p when she seems to see that p is not a bad disposition for a subject to have in the least: most of the time when subjects seems to see that p , they really do see that p , and p is true. It is only rarely the case that one merely seems to see that p . There is not better set of dispositions to be had.

By contrast, take the method telling one to believe p when one sees that p in the presence of evidence for thinking that one's senses are not to be trusted. A subject who adopts this method is also disposed to believe p when she merely seems to see that p in the presence of evidence for thinking that her senses are not to be trusted. But then, her overall dispositions could be significantly improved given the goal of knowledge, or even of true belief. For a subject who follows the rule in question is disposed to stick to a belief in p whenever she seems to see that p and there is evidence for thinking that her senses are not to be trusted. Of possible broadly normal cases in which such a disposition might be manifested, a significant portion are ones in which the evidence about one's senses is *not* misleading and hence, had better be taken seriously. After all, it's not as if evidence of a defeating kind tends to be more

²³ Since they make appeal to the dispositions and abilities that accompany the adoption of a method, such evaluations depend on the cognitive architecture of the subject under consideration. For instance, a possible subject with a very powerful ability to intuit, in an immediate manner, all mathematical truths, does well to follow a rule telling her to believe a mathematical truth just in case it seems immediately obvious to her, but such a rule would clearly not be a very effective way for actual subjects to acquire mathematical knowledge.

²⁴ I do not here intend to be discussing a rule telling one to believe p when one sees that p no matter what other evidence one has. One might instead want to characterise the intended rule as one telling a subject to believe p when she sees that p and there is no evidence for thinking that her senses are not to be trusted.

²⁵ Why not then say that the subject has just one disposition, the disposition to believe p when she seems to see that p ? And then, why not say that she follows just one rule, the rule telling her to believe p when she seems to see that p ? I suspect the motivation for the proposed move to be the assumption that while a subject can be mistaken about whether she sees that p , she can never be mistaken about whether she seems to see that p . But there are strong reasons to think that this assumption is simply false: why couldn't it merely seem to a subject as if it seems to her as if p ? Then, a disposition to believe p when one seems to see that p would be accompanied by a disposition to believe p when it seems to one as if one seems to see that p . I fear the dispositions sought by proponents of this move are simply not to be had. See, in particular, Williamson's (2000, Ch 8) argument against the non-phenomenal individuation of rules.

misleading than evidence that doesn't play a defeating role – indeed, this is why it should be taken seriously. It would be much better to be disposed to suspend judgment in such cases. This is why the rule *believe p when you see that p in the presence of evidence for thinking that your senses are not to be trusted* is not part of a policy that is knowledge conducive in the intended sense. A reasonable subject would not adopt or follow such a rule, even though it is success entailing.

Hence, some methods that are perfectly good in the sense described above nevertheless don't belong to policies that are knowledge conducive for beings anything like ourselves. Such methods are not good methods to adopt. This applies, in particular, to methods recommending dogmatist adherence to previous beliefs in the face of new evidence. Subjects who follow policies recommending them to retain knowledge-constituting beliefs in defeat cases are susceptible to practicing dogmatism in “bad” cases in which the new evidence is not at all misleading. In these cases it will seem to them as if they are following the same method as in good cases, thereby retaining knowledge, whereas they will in reality be retaining beliefs in falsehoods. Overall, it would be much better to follow a policy recommending the revision of belief in the light of new evidence. This is why reasonable subjects adjust their beliefs in defeat cases.

It is important to note that on the kind of view sketched, subjects who retain knowledge in defeat cases are genuinely criticisable. After all, they fail to follow policies that it is rational for someone with the goal of acquiring knowledge, or even just of true belief, to adopt. They manifest dispositions that are bad from the perspective of this goal, dispositions that, overall, cannot be expected to do a good job on occasions on which they are manifested. It is worth emphasising that there is something rather deviant about knowing while manifesting a disposition that will often give rise to false belief. Moreover, this deviance goes beyond the deviance of believing truly while manifesting such a disposition: after all, even a completely random method can have some non-zero chance of producing a belief in a truth, but not so for producing a belief that could not easily have been false. Generally, to form a belief that is true in the sort of non-accidental manner required for knowledge while manifesting bad dispositions, a subject has to be in special circumstances shielding her belief from error.

6. Reasonableness and paradigm cases of knowledge

My strategy for explaining away intuitions about defeat appeals to a tendency to think that when a subject is not acting reasonably, she lacks knowledge. Why might we have such a tendency?

I pointed out above that circumstances in which one knows while being unreasonable are deviant. Reasonable belief is a feature of ordinary, paradigm cases of knowledge, cases through which we learn to apply the term. What bears emphasis is that in such cases, reasonableness is not merely a superficial manifestation of the underlying nature of knowledge. For typically, the reasonableness of a belief enters into an explanation of why and how a subject manages to know.²⁶ In particular, it

²⁶ In this respect, the relation between being reasonable and knowing is different from, for instance, the relation between the superficial features of water and water. In paradigm examples of water, it's not as if the superficial features explain why the underlying nature is present. It is rather the other way around: the deeper essence explains why it is that samples of water in specific sorts of environments have the superficial features they do.

explains why and how she manages to form a belief that is not merely true by luck, or true in an accidental way. For example, when I form an ordinary perceptual belief about the weather, I practice certain competences, and it is because of those competences that I form a belief that is not only true, but true in a non-accidental way. I come to believe that it is sunny, but had it been cloudy, my perceptual abilities would have picked up on this, and I would have come to believe that it is cloudy. I have knowledge not because I am in circumstances shielding me from error, but because I practice a general ability enabling me to do so myself. In paradigm cases, there is a sense in which having knowledge depends on the belief in question having been formed in a reasonable manner. As a consequence, asking whether a belief is reasonable is a useful heuristic for evaluating whether it counts as knowledge.

It may also be worth mentioning an additional, related source of our reluctance to attribute knowledge in defeat cases, which has to do with the fact that attributions of knowledge are acknowledgements that a subject has done something right: it's not just that the content of her belief is true, but the belief was formed in a way that resulted in believing the truth in a non-accidental way. But we are rightly reluctant to positively evaluate the methods employed by subjects in defeat cases, and this may partly explain why we are reluctant to attribute knowledge. Perhaps we don't want to reward the adoption of a method that is not part of a knowledge conducive policy. But moreover, it is reasonable to expect positive acknowledgements to reinforce and even create dispositions, thus encouraging similar behaviour in the future. And such encouragement can be expected to have unhappy consequences: attempts to replicate what she does when retaining knowledge in a given defeat case are likely to result in retaining false beliefs on other occasions.

I have sketched a view on which it is possible to know despite practicing defective strategies for acquiring knowledge, and despite failing to manifest general dispositions to know. But it is worth pointing out that what emerges from the discussion is an externalist view that is an alternative to the one I have defended. For given that my account of reasonableness is thoroughly externalist, why not opt for a theory that adopts something very much like reasonableness as a condition on knowledge? Why not say, roughly, that for some x , knowledge is x (such as safe belief, or even just true belief) achieved through a general disposition to achieve x ? Such a view would come at least very close to virtue theoretic approaches put forth in the literature.²⁷ I cannot here go into considerations of naturalness or simplicity, even though I don't think they should be given short shrift especially by anyone who thinks of knowledge as a natural epistemic kind. But let me return to a general line of thought already mentioned above.

The epistemology I have sketched relies in an important way on a distinction between two different kinds of assessments: assessments of how well things go on a particular occasion, and assessments of how well a subject is disposed to do more generally, across a wide range of somewhat normal cases. We draw such distinctions in numerous domains, and it is clear why we need to do so. Too much emphasis on dispositions that achieve the desired result across a wide range of cases leads to evaluative notions that are too difficult to satisfy. It comes as no surprise that it is very difficult to think of things for which the correct theory takes the form, for some x , of *achieved through a general disposition to achieve x* .

Despite this, one might object that knowledge is special: it cannot be achieved while failing to manifest good dispositions. However, even those who emphasise the

²⁷ See, for instance, Ernest Sosa (2007) and John Greco (2009).

role of competences and good dispositions in their theories need the distinction between how well things go on a particular occasion and how well a subject is disposed to do more generally. For instance, Sosa defends a view on which (animal) knowledge is belief that is true because competent, where competence is at least very closely connected with a disposition to form true beliefs. However, he admits that an archer's shot made from a long distance can be competent and skilful despite the fact that a gust of wind was required for carrying the arrow all the way to its target.²⁸ In such cases, the archer's dispositions only succeed in cooperative environments. Similarly, sometimes knowing requires cooperation from one's environment. But this is already a step away from the kind of view under discussion on which for some x , knowledge is x achieved through a general disposition to have x . Once this first step is taken, I am urging we go all the way and stop thinking of knowledge as being special. Like most other things, it can sometimes be achieved through bad strategies.

7. Concluding remark

Some things are hard to achieve because it is very difficult to acquire the right strategies – including the right dispositions and abilities – enabling one to achieve them. But other things are hard to achieve because they are not achieved by following what is generally good practice. Instead, one must employ methods that don't make up policies that it is good to adopt, or manifest dispositions that are very likely to get one into trouble on other occasions.

For instance, it may be that given how difficult a given route up Annapurna is, and how many unpredictable factors success depends on, no strategy that is successful in a single case is going to be very effective in getting to the top in a wide range of cases. A mountaineer may achieve a first ascent by carrying no extra gas and turning back at a time that is very late. But perhaps the route is so difficult that it can only be climbed by daredevils following bad practice: carrying no extra gas enters in an important way into an explanation of the mountaineer's achievement, for had her bag been any heavier, she would not have been able to climb a difficult technical bit. Despite her achievement, we are right to criticise the mountaineer. Given that she wants to climb many mountains, she was not acting in a reasonable manner. Similarly, on the view I am defending, there are situations in which knowledge is difficult to achieve because holding a belief in those situations requires following bad practice. I have argued that defeat cases are a case in point. But in the epistemic domain as in others, foolishness is sometimes rewarded.²⁹

References

Armstrong, D. M.
1973 *Belief, Truth and Knowledge*, Cambridge: Cambridge University Press.

Bergmann, M.

²⁸ Sosa (2007).

²⁹ I am grateful to Tim Williamson, Ernie Sosa, John Hawthorne, Ralph Wedgwood, Scott Sturgeon, Daniel Morgan, Troy Cross, Ville Aarnio, and audiences at the University of Michigan, University of Toronto, and University of Oxford.

- 2000 'Deontology and Defeat', *Philosophy and Phenomenological Research*, 60:1, 87-102.
- 2006 *Justification Without Awareness*, Oxford: Clarendon Press.
- Boghossian, P. A.
 2008 'Epistemic Rules', in *Content and Justification*, Oxford: Oxford University Press.
- Bonjour, L.
 1980 'Externalist Theories of Empirical Knowledge', *Midwest Studies in Philosophy*, Vol V, pp. 53-73.
- Chisholm, R.
 1966 *Theory of Knowledge*, Englewood Cliffs, N.J.: Prentice-Hall.
- Goldman, A. I.
 1986 *Epistemology and Cognition*, Cambridge, MA: Harvard University Press.
- Greco, J.
 2009 'Knowledge and success from ability', *Philosophical Studies*, 142 (1).
- Nozick, R.
 1981 *Philosophical Explanations*, Cambridge, MA: Harvard University Press.
- Pappas, G. S. & Swain, M. (eds.)
 1978 *Essays on Knowledge and Justification*, London: Cornell University Press.
- Pappas, G. S. & Swain, M.
 1978 'Introduction', in G. S. Pappas & M. Swain (eds.) 1978, pp. 11-40.
- Pollock, J. L. & Cruz, J.
 1999 *Contemporary Theories of Knowledge*, 2nd edition, Lanham, MD: Rowman & Littlefield.
- Pollock, J. L.
 1995 *Cognitive Carpentry: A Blueprint for How to Build a Person*, Cambridge, MA: MIT Press
- Sosa, E.
 1999a 'How to Defeat Opposition to Moore', in Tomberlin (ed.) *Philosophical Perspectives* 13, pp. 141-153.
 1999b 'How Must Knowledge Be Modally Related to What Is Known?', *Philosophical Topics* 26, pp. 373-384.
 2007 *A Virtue Epistemology: Apt Belief and Reflective Knowledge*, Volume 1, Oxford: Oxford University Press.
- Wedgwood, R.
 2002 'The Aim of Belief', *Philosophical Perspectives* 16, pp. 267-297.
- Williamson, T.

- 2000 *Knowledge and its Limits*, Oxford: Oxford University Press.
- 2009 'Reply to Hawthorne and Lasonen-Aarnio', in P. Greenough and D. Pritchard (eds.), *Williamson on Knowledge*, Oxford: Oxford University Press.