

## Losing knowledge by thinking about thinking

Penultimate draft. Final version forthcoming in *Reasons, Justification and Defeat*, Jessica Brown and Mona Simion, eds., Oxford University Press.

**ABSTRACT:** Defeat cases are often taken to show that even the most securely-based judgment can be rationally undermined by misleading evidence. Starting with some best-case scenario for perceptual knowledge, for example, it is possible to undermine the subject's confidence in her sensory faculties until it becomes unreasonable for her to persist in her belief. Some have taken such cases to indicate that any basis for knowledge is rationally defeasible; others have argued that there can be unreasonable knowledge. I argue that defeat cases really involve not an exposure of weakness in the basis of a judgment, but a shift in that basis. For example, when threatening doubts are raised about whether conditions are favorable for perception, one shifts from a basis of unreflective perceptual judgment to a basis of conscious inference. In these cases, the basis of one's knowledge is lost, rather than rationally undermined. This approach to defeat clears the path for a new way to defend infallibilism in epistemology, and a new understanding of what can count as the basis of any instance of knowledge.

Knowledge can be real without being armour-plated. We gain and lose knowledge in many ways in the normal course of life. We learn and then naturally forget, within minutes or hours, many details of what we have perceived for ourselves and heard from others. We can also lose knowledge without losing belief; for example, when the world changes in unseen and unexpected ways behind our backs. In addition to passive processes of forgetting and becoming outdated, we can actively switch to suspension of judgment on points where we once had knowledge. This last type of switching is often routine and unremarkable; for example, when one decides to suspend belief upon recognizing that one's grounds or basis for judgment on a given point may well have weakened over time. Upon reflection, I have doubts about the accuracy of my mental map of the city where I used to live two decades ago, a city I used to know quite well. The quality of my judgments about Albuquerque has suffered from erosion on two fronts: the city's businesses, streets and bus routes have doubtless changed over time, and my memories have been shifting and fading. Given this natural deterioration of my basis, it is entirely reasonable—and epistemologically uncontroversial—for me to suspend judgment on any finer points of that city's layout. My current state of suspension casts no shade on the epistemic status of the judgments I routinely made in this area twenty years ago.

Other cases of switching into suspension seem potentially more troubling. To some, these cases suggest a certain tension between rationality and knowledge, or some rational fragility in the bases of the judgments which constitute knowledge. Known in the literature as “defeat cases”, these scenarios are sometimes taken to show that it is possible to have knowledge of a proposition on a

certain basis, and then find oneself rationally required to switch to suspension of judgment, despite continuing to have exactly one's original basis for judgment. Others have argued that even if these cases show that it is in some sense unreasonable to cling to the originally knowledge-sustaining basis of one's judgment, one can nonetheless retain knowledge by doing so.

This chapter argues that defeat cases have been misinterpreted by both sides. Knowledge is indeed lost in defeat cases, but it is not lost while one's original justifying basis is retained: a deep feature of these cases is that they inevitably involve epistemically significant—although sometimes hard-to-spot—shifts in the basis of one's judgment of the original first-order question. These shifts are produced in various ways, including the production of self-conscious worries about how one is thinking. The new basis that arises after the shift will not sustain knowledge of the key proposition, but it is not the same as the basis one had before the self-conscious worries or other defeating conditions arose. Close examination of these cases can help us gain a sharper sense of what actually constitutes the basis of a judgment, and why exactly the bases that underpin knowledgeable judgments are strictly immune to rational undermining, while nevertheless being vulnerable to loss through what I will call basis displacement. The fact that knowledge of a given proposition can be lost does not show any weakness in its original basis: just as you might come to forget ever having felt the pain of some particular toothache, where this forgetting casts no shadows over the epistemic credentials of your originally vivid knowledge of the pain, so also you can lose knowledge through the displacement of your original basis for judgment, notwithstanding the optimal epistemic quality of that original basis.

Section 1 outlines my infallibilist approach to justification and the more popular fallibilist alternative. Section 2 aims to explain the appeal of fallibilism, and Section 3 examines defeat cases in the light of my theory of justification.

### **1: Justification and the basis of a judgment**

Despite immense debate over just what is meant by 'justification' in epistemology, there is widespread agreement that it denotes a positive normative status, and that this status applies at least to all states of knowledge. My approach here will be to start with the justification of what is known, before moving outward to more controversial cases.

What is the justifying basis of a person's knowledge of a fact? We might start by observing that facts (or true propositions) can be known in a variety of ways. It is possible to know a fact—say, where some object is currently located—through several distinct sensory modalities, through appropriate testimony from a trusted knower, through sound inference from premises themselves known through observation or testimony, or through some combination of these. In addition to the outer-world channels of perception, inference, and testimony, we also have interoceptive awareness of facts about inner phenomenal states such as thirst and dizziness. The range of ways in which propositions may be believed includes all these ways of knowing—given that knowing entails believing—plus a wide variety of epistemically worse ways of coming to make a judgment. One might come to believe a proposition on the basis of an illusion, deceptive testimony or unsound inference, or in some other way, perhaps even as a result of a novel medical intervention. One might wonder what unites the more restricted set of ways in which propositions can be known, as opposed to just believed. If we take knowing to be distinguished from believing by its special relationship with the truth—it is an essential feature of knowing that only what is true can be known—then ways of knowing are likewise most naturally distinguished from the broader class of ways of believing in virtue of their relationship to the truth. In my view, a way of judging that  $p$  is a way of knowing only if the truth of  $p$  is essential to this way of judging.<sup>1</sup> On this way of individuating “ways of knowing”, *inference* is not a way of knowing, but *sound inference from known premises* is. Likewise, *perception* counts as a way of knowing only if it is understood in success terms, as involving sensory systems that have an appropriate causal sensitivity to a range of facts, where some fact within this range causes a corresponding judgment through the normal operation of these systems. Although “perception” is sometimes used in a broader sense that encompasses any kind of sensory activation, I will use it in the more restricted success sense here. Similarly, *testimony* will denote acceptance of an honest knower's word, rather than the absorption of any report. With this restricted understanding of ways of knowing in hand, my proposal here is that the justifying basis of a person's knowledge of a fact—what makes that attitude epistemically right, indeed optimal—is nothing other than the way it is known. While my initial focus is on the justifying bases of knowledge, I do not mean to imply that judgments falling short of knowledge can have no justifying bases whatsoever; their weaker, derivative level of justification is a function of how closely the ways in which they were formed resemble ways of knowing (further details will emerge in due course).

---

<sup>1</sup> Note that this is a necessary, and not a sufficient condition; it may be essential to *guessing-correctly-on-a-Tuesday* that the relevant proposition is true, but this will not be enough to qualify it as a way of knowing.

This basic approach to the relationship between knowledge and justification has old roots in the history of philosophy, most vividly in the Classical Indian Nyāya tradition. In the Nyāya Sutra (c. 200CE), Gautama characterizes perception as “inerrant”(1.1.4), and stipulates that it must involve a causal connection between the perceived object and the sensory faculty (Dasti and Phillips 2017). Other ways of knowing (*pramāṇas*, in Sanskrit) such as testimony and inference, are likewise described in the tradition as “faultless”, “unfailing”, and “non-deviating” (Matilal 1986, 135). Less-than-sound inference is labelled “pseudo-inference”, hallucination and the like are “pseudo-perception”, and misinformation is not testimony but “pseudo-testimony” (Dasti and Phillips 2010, Dasti 2012). Contemporary non-skeptical epistemologists will readily agree that the good cases here have some justifying basis; what is more controversial is to identify the justifying basis of a known fact with the way it is known.

This understanding of the justifying basis of what is known is frankly incompatible with the currently popular fallibilist approach to knowledge, according to which there is typically a significant gap between one’s knowledge of a proposition and one’s justifying basis, so that what is known can have the same justifying basis as what is merely believed, or even falsely believed. Here is Baron Reed’s gloss on the core idea of fallibilism:

Roughly stated, the basic idea is that the subject can know something even though it *could have been false*. This is not the same as saying that the subject can know something that is false – it is very widely accepted by philosophers that, if a belief counts as knowledge, it is true. Rather, the claim is that a belief held with a particular epistemic grounding can be knowledge even though the subject could have held that belief *with the same grounding* in circumstances where the belief is false (and, of course, in those circumstances the belief would not count as knowledge). (Reed 2012, 585)

According to fallibilists, false beliefs can have precisely the same epistemic grounding as instances of knowledge, so epistemic grounding in cases where *p* is known cannot be understood in a way which makes the truth of *p* essential to one’s epistemic grounding for *p*.

Not all formulations of fallibilism are given in terms of epistemic grounding; some speak of evidence or justification. Another standard formulation of fallibilism is rendered by Reed as follows: “A person S knows that *p* in a fallibilist way just in case S knows that *p* on the basis of some justification

*j* and yet *S*'s belief that *p* on the basis of *j* could have been false (or mistaken or in error)."<sup>2</sup> (2012, 586) The example he offers to bolster his point will bring out the difference between our views of justifying bases:

For example, I know that the Cubs beat the Dodgers the last time they played; I know it because my brother told me what happened, and he is usually reliable about this sort of thing. But, if he had misread the box score in the newspaper, I still would have believed him. In that case, my belief would have been held with the same justification, but the belief would have been false. Because this case is possible, the knowledge I actually have (when my brother did not misread the box score) is fallible. (2012, 586)

For simplicity, I will understand the subject of these cases as hearing from his brother just that the Cubs won, without any further elaboration about his having read the paper as opposed to having watched the game live or on television. I will also understand the subject to be taking his brother's word for it, in both cases, as opposed to forming his belief on the basis of explicit reasoning from the premises, "My brother said that the Cubs won," and "My brother is usually reliable about this sort of thing." We generally accept trusted testimony without such calculated reasoning about its likelihood of being correct, and if we were to restrict ourselves to such explicit reasoning, it is not clear that we would gain as much as we do from the reports of others. If explicit reasoning from exactly the premises just given yields knowledge, it seems only to yield knowledge of a weaker conclusion like, "It is highly likely that the Cubs beat the Dodgers", as opposed to knowledge of the stronger conclusion that the Cubs won. I take Reed's use of the stronger conclusion in this example to indicate that he also accepts trusted testimony as a path to knowledge that is interestingly different from the explicitly reasoned calculation that could deliver knowledge of the weaker claim. Indeed, the weaker claim (that it was likely the Cubs had won, given the report of a typically reliable person), could be true, and arguably even known, even in the case of the misread loss, so if the contrast between the cases is to illustrate that the very proposition known might have been false, we should not weaken it.

---

<sup>2</sup> In the face of concerns about knowledge of necessary truths, this provisional formulation is later amended to "S knows that *p* in a fallible way just in case *S* knows that *p* on the basis of some justification *j* and yet *S*'s belief that *p* on the basis of *j* could have failed to be knowledge." (Reed 2012, 587) This amended formulation remains incompatible with my account.

I agree with fallibilists like Reed that knowledge can be transmitted through testimony, even third hand, as from a brother's report of a newspaper story. However, the view I will defend distinguishes different justifying bases for the good case and the bad: in the good case, I take the justifying basis to consist in the entire conduit from the Cubs' actual victory to the subject's awareness of that victory, incorporating the successful perceptual and inferential processes giving rise to knowledge on the part of the reporter, and the safe transmission of this knowledge to the brother who reports what he came to know by reading the paper to the appropriately trusting subject. It is because the subject has learned in such a way from a knowledgeable source that his judgement is epistemically in the clear, normatively appropriate, fully justified. In the bad case, there is no such chain of transmission between the outcome of the game (presumably a loss) and the subject's formation of a belief in a victory. Misreading a report is not a way of knowing, nor is trusting someone who is mistaken. There are some commonalities between the good and bad cases: for example, in both, the subject trusts someone who is usually reliable. Thanks to the partial similarities between the knower and the subject of the bad case, even the bad case subject has some measure of justification (on which more later), but the justifying bases of the good and the bad cases are not the same.

On the common fallibilist understanding of these cases, by contrast, the subjects of the good and bad cases can have precisely the same justification.<sup>3</sup> One way to make sense of this would be by identifying justification with what is available to the subject to report as the basis of his judgment that  $p$ , if asked how he knows that  $p$ . What matters here is not the actual history of the subject's judgment, but only those aspects of the apparent history which are readily available to the subject for report. This subjective impression of one's basis will be called the subjective basis, for short,

---

<sup>3</sup> I say "can have" rather than "always have", to admit a greater variety of types of fallibilism. For some forms of internalist fallibilism, the subjects of good and bad cases like the ones just given will always have exactly the same justification; there is no gap between what one's justifying basis is and what it seems to be (e.g. Smithies 2019). But even for externalist fallibilists like Goldman (1979), who would assign sharply different levels of justification to the perceiver and the hallucinator, running distinct cognitive processes with different levels of reliability, there are at least some phenomenally indistinguishable pairs of good and bad cases where the same process is at work across the divide. For example, if the same retinal stimulation is produced by a convex object lit from above and a concave object lit from below, generally resulting in a perceptual impression of convexity, given our learned background expectations of light from overhead, then the same reliable but not infallible process is at work in the person who sees the normally-lit convex object as convex, and the person who sees the concave object as convex in the presence of subtle lighting from below. These good and bad cases are equally justified, in Goldman's later fallibilist reliabilism (although not in his earlier infallibilist causal view), given his restriction on the individuation of belief-forming processes to what happens "within the organism's nervous system" (1979, 12). Although for simplicity I focus largely on internalist forms of fallibilism here, I believe there are ways of extending the arguments of this chapter to positions like this kind of reliabilism.

although we should be careful to note that the subjective basis is not a basis for the judgment in the causal sense of that word, even if it represents some elements in the causal basis. Between this subjective fallibilism and my objective infallibilism lie a spectrum of views, including externalist forms of fallibilism such as process reliabilism; for simplicity, I will focus mainly on the extreme ends of the spectrum in what follows.

We have a limited view of the causal processes behind our judgments, even in cases where we gain knowledge. The subject who comes to know about the victory can trust his brother without being able to report on the exact track record supporting his sense that his brother is reliable about “this sort of thing” (and indeed without being able to explain the exact scope of “this sort of thing”), although there may well be some intelligent appreciation of the brother’s track record underlying this trust. Consciousness presents products of cognition, rather than the underlying processes, let alone the distal features of the processes that secure our grip on external truths. It is because what is available to consciousness is so restricted that the bad case (pseudo-testimony) can be mistaken for the good (testimony): from the subject’s first-person perspective, I can equally respond to a challenge in either case by saying something like, “my brother said that the Cubs won, and he is usually reliable about this sort of thing.”

In my view, the similarity of reports produced in the good and bad cases is evidence of the disparity between subjective and objective bases: one might reasonably say the same thing about one’s basis in a case of error and in a case of knowledge, and in either case, one has limited first-person access to what actually makes one’s judgment normatively appropriate. The fact that one’s judgment is normatively appropriate in the good case means that one does have suitable first-person access to the key first-order fact (for example, that the Cubs won), but it is possible to know that the Cubs won without knowing the full story behind how it is that you know this fact, down to the relevant characteristics of the reporter who saw the game, and so forth. Despite failing to capture the full story of how we are normatively in the clear, subjective impressions of our ways of thinking nevertheless play an interesting role in our cognitive economy, and are not simply epiphenomenal to the course of our thinking about first-order issues over time. In particular, conscious representations of the bases of our judgments play a vital role in our cognitive flexibility, our capacity to change how we are thinking about first-order issues, whether such changes are prompted by anomalies we encounter as individuals, or by challenges we receive from others.

## 2. The function of subjective representations of basis

The appeal of fallibilism can be better understood if we undertake a deeper investigation of the character and function of conscious representations of the bases of our judgments. One intriguing feature of these conscious representations is their pervasive availability. While our attention is ordinarily directed at first-order questions in the world—one feels elation at the Cubs’ victory, one notices a large object in the room—these first-order questions are for self-conscious creatures like us always settled with some accompanying representation of how they are settled (through testimony, through visual or haptic perception, and so on). There are consciously available differences between seeing or touching an object, and between judging that the Cubs won on the basis of having seen them play, having inferred that they must have won, or having heard about it via testimony. It is a good question why we have this kind of metacognitive awareness, and what relation it ultimately bears to the normative status of our judgments.

To answer these questions, we can begin by reviewing some evidence bearing on the general availability of subjective representations of basis. In roughly a quarter of the world’s languages, ordinary declarative judgments bear obligatory markings of the distinctions between major ways of knowing such as first-hand perception, testimony, and inference (Aikhenvald 2004), but the source monitoring functions that underpin those grammatical markings are taken to operate in typical adult speakers of all languages (Johnson, Hashtroudi et al. 1993, Papafragou, Li et al. 2007). In earlier work, I emphasized the social function of source monitoring: if we have some indication of where our judgments are coming from, we can do a better job of sharing our knowledge with others, and where needed, selectively telling them what they don’t already know (Nagel 2015). Source monitoring also enables us to share our perceptual knowledge of features of our current context by directing others’ perceptual attention appropriately (“look over there!”), sparking their own independent way of knowing the fact in question. Another function of source monitoring is connected to our individual cognitive flexibility: consciously available representations of the bases of our judgments guide us in our active epistemic exploration of the world (stepping closer, turning the lights on, moving our hands or gaze over the object), and in rethinking matters when something we have judged to be the case turns out to be problematic. Realizing I have taken someone else’s word for a disputed claim, I can try to verify it with my own eyes; if what I seem to see strikes me as implausible I can ask others if they also see it. These tasks of sharing and rethinking can be

combined: in epistemically cooperative communities, the work of sorting testimony from pseudo-testimony (and so on) is socially distributed. Others are often better positioned to tell whether you are seeing or hallucinating, or whether your source is trustworthy. So, even in languages such as English, where source marking is lexical and optional, we frequently communicate source, marking what is inferred with a construction like “I gather that..” or what is derived from testimony with “I hear that..”. These constructions can be used to invite those who may be better positioned epistemically to corroborate or correct our judgments (Pomerantz 1980).

The social role of subjective bases does something to explain the natural appeal of fallibilist views of justification. Even if there is something normatively undesirable about cases of pseudo-perception or pseudo-testimony (other things being equal, surely these are not the conditions we hope to inhabit), the subjects who find themselves in these conditions will behave just like their good case counterparts when their epistemic status is challenged. They will tell similar (partial) stories about how they arrived at their judgments, including to themselves, if they become worried about their own epistemic standing and attempt to introspect. As far as their active control of their epistemic agency is concerned, subjects are guided by their subjective bases, so theories that take justification to be a matter of the proper exercise of epistemic agency will place considerable weight on subjective bases.

Explicitly inferential reasoning deserves special attention in this account. Reasoning might seem to be an exception to my contention that the subjective impression of one’s basis is only a partial representation of what makes the judgment normatively appropriate. In particular, when one is conscious of the premises of a deductively sound argument, and then derives the conclusion, one might think that the subjective and objective bases of the judgment coincide, and that in this type of case, the subjective basis fully captures what causes the judgment and makes it normatively appropriate. However, even here there are many normatively relevant features of the inferential judgment that are not presented to consciousness. If one is to attain knowledge of the conclusion, it matters not only which premises manifest themselves in consciousness (say, by being rehearsed in inner speech), but also the epistemic standing of these premises (which will typically depend on factors not present in consciousness), and the soundness of the rules of inference sparked by the recitation of those premises to produce the conclusion. Explicit reasoning, like all reflective cognition, consists in operations performed on consciously available material, but most of the

workings of reflective cognition are themselves unconscious (Evans 2009). We can fail to notice this because typical humans share common capacities to deploy common rules of inference, including for example simple rules of syllogistic reasoning (Mercier 2011, Mercier and Sperber 2011), so in cases where the epistemic standing of one's premises is not in question, one can share inferential knowledge of a conclusion by reciting one's premises in a way that will equally trigger the subconscious operation of the same relevant rules of inference in oneself and one's audience, to secure the transition to one's conclusion. Thanks to the typical commonalities in our rational capacities, explicit reasoning is a powerful device for the sharing of knowledge: if my epistemic standing on a given point is initially in doubt, but I can produce a sound demonstration of that point from premises known to my audience, reliant on rules of inference they also deploy, then what I transmit to the audience is not only knowledge of the conclusion, but also the activation of a way of knowing it. They do not need to trust me, or see me as knowing, to gain knowledge of the conclusion we arrive at together. This is one reason why we are so naturally inclined towards explicit argumentation when we encounter resistance from an audience, or doubts about our epistemic standing.

In the face of doubts about our epistemic standing on a given point, we can engage in explicit reasoning, even if our original judgment on this point was not made on the basis of reasoning. If I have seen a highly unusual event outside the deictic sphere of our current conversation, I cannot share my knowledge of the event by directing my audience's perceptual attention to it. If I attempt to transmit my knowledge of the unusual event through simple testimony, I face the risk of being dismissed as failing to know, given that what I am reporting will sound unlikely. A better strategy will be to present whatever evidence I can share to support my claim that the event occurred, pointing out its causal traces, and so forth, even though my own awareness of the event was secured through first-hand perception and not inference. Where a truth is itself known through explicit reasoning, there is little difference between the strength of one's dialectical position and the strength of one's epistemic position; in other cases, dialectical strength is used as a test for the presence of knowledge, and there may well be a gap between whether one can pass this social test for knowledge and whether one actually knows.

There are internalists who make inference essential to justification, seeing even perception as essentially involving inference from prior evidence. For example, Stephen Schiffer describes the

formation of a perceptual belief (say, that there is a red cube before him, abbreviated ‘Cube’) in a favorable case (‘Good’) as follows: ‘in Good I come to be justified in believing Cube by inferring it from the fact that I am having such-and-such sensory experiences as of Cube, and I cannot become justified in believing Cube in Good other than by inferring it from that evidence’ (2009, 198). This position is problematic in several ways. One might worry that such an inference would not actually be a very good one: from the fact that I am having sensory experiences “as of” the presence of a cube, it does not actually follow that there is a cube in front of me. Such experiences might be caused by hallucinations or holograms as well as cubes. One might worry about one’s entitlement to the “as of” locution, which seems to presuppose some prior acquaintance of the experiences that correlate with the real presence of cube, where knowledge of this correlation will be hard to secure. And lastly, one might worry about the reality of such inferences: if we take it to be a condition on “evidence” that it is evident in the sense of being present to consciousness, ruling out accounts in which the relevant inference is subpersonal, then it seems that perceptual beliefs on Schiffer’s account will only very rarely count as epistemically justified. How often are we first conscious of having a certain type of “as-if” seeming, and then only subsequently aware of the outer object, through inference?

By contrast, externalists like Timothy Williamson do not insist that justified belief is ultimately inferential in character, and do not see basing as necessarily an explicitly inferential relation. Alongside the well-recognized category of explicitly evidence-based belief, Williamson introduces a category of implicitly evidence-based belief, where a belief in  $p$  is in this category “if it is appropriately causally sensitive to the evidence for  $p$ ” (Williamson 2000, 191). Evidence itself, meanwhile, is not a precursor to knowledge but knowledge itself, and knowledge does not in general need to be based on prior evidence (or knowledge). Like many other externalists, he emphasizes the naturally unreflective, and non-inferential character of perception: “When we acquire new evidence in perception, we do not first acquire unknown evidence and then somehow base knowledge on it later. Rather, acquiring new evidence *is* acquiring new knowledge. That knowledge need not itself be based on further evidence, nor is it evidence for itself in some non-trivial way. But it is evidence for or against potential answers to questions to which we do not yet know the answer.” (Williamson 2014, 4)<sup>4</sup> On this approach, which fits and indeed inspires the view I am defending in this chapter,

---

<sup>4</sup> One might wonder whether this more recent formulation of Williamson’s view is a strengthening of the earlier view, according to which a perceptual belief that  $p$  could be implicitly evidence-based by being appropriately

we are not initially aware of visual appearances as of the cube, from which some belief about the cube is then inferred; in the good case, what we are initially conscious of is the cube. Our first moment of evidence is that there is a cube before us, in the good case. If we turn introspective, directing attention inward, we are aware that we have specifically visual experience “as of a cube”, but this is a metacognitive byproduct of the original unreflective judgment, and not a prior condition of it. It is true that hallucination would generate a similar metacognitive byproduct, but false that hallucination would make us equally justified. In the bad case, all we are conscious of is that there is something which appears to be a cube, but it is part of the badness of the bad case that we will typically judge that there is a cube present. If we wish to make a judgment which will be legitimate in both cases, we can restrain ourselves to the more cautious judgment that a cube seems to be present, but a general strategy of restricting oneself to such judgments would entirely undermine the value of our world-sensitive perceptual mechanisms. It is a feature of unreflective cognition that it does not involve reflection on how we ought to be thinking: in unreflective cognition such as ordinary perceptual judgment we are focused on the world, and not on the question of how we should now be making judgments.

Because metacognitive impressions of our judgments play a key role in defending our epistemic status under challenge, it is not easy to give up the idea that instances of knowledge always require the backing of prior evidence. In her defense of fallibilism, Jessica Brown observes that there is something quite counter-intuitive about positions like the one I am defending here. She defines both fallibilism and infallibilism in terms of a relation obtaining between evidence and judgment: what infallibilists hold and fallibilists deny is that, “knowing that  $p$  requires evidence which entails that  $p$ ” (Brown 2018, 4). She observes that Williamson’s equation of knowledge with evidence (in Williamson 2000) gives rise to an unusual form of non-skeptical infallibilism: because any known proposition is automatically included in one’s evidence set, the fact that every proposition entails

---

causally sensitive to the perceptual evidence for  $p$ . On the old position, one could have perceptual evidence for  $p$  without knowing that  $p$ , for example if one had misleading evidence about the source of light. However, even in the new position, which highlights situations in which we just gain the perceptual evidence *that*  $p$ , thereby coming to know that  $p$ , we can still distinguish a separate notion of “perceptual evidence *for*  $p$ ”, where the knowledge we gain speaks in favour of  $p$  without entailing knowledge of  $p$ . This kind of knowledge might be especially important in cases where we have misleading evidence speaking against  $p$  which makes us think inferentially about the question of whether  $p$ , rather than forming a naïve perceptual belief. Thanks to Carlotta Pavese for valuable discussion on this point.

itself will ensure that no proposition is held without entailing evidence (Brown 2018, 32). She then argues that it is inappropriate to cite a proposition as evidence for itself (2018, ch.3).

Where the function of evidence is to lend support to a contested proposition, citing that very proposition will beg the question at issue, whether the contest arises from some audience who could have doubts about whether the speaker knows that  $p$ , or whether it arises from the speaker's own doubts about her epistemic status with respect to  $p$ . So, Brown is right that one's knowledge that  $p$  is generally not self-supporting in a way which will satisfy those who have raised doubts about one's knowledge that  $p$ . However, if we can separate the task of proving one's epistemic status from the conditions of initially entering that status, we do not need to insist that every item of knowledge be backed by prior evidence. If our sensory systems give us appropriate causal sensitivity to some features of reality, for example, then when these systems make us aware of those features, we can gain knowledge without prior supporting evidence.

Brown advances two independent lines of argument against the kind of strategy I advance at the outset of this chapter, in which ways of knowing constitute our sources of justification. If I insist that a way of judging that  $p$  only yields full justification when the truth of  $p$  is essential to one's judgment that  $p$ , Brown objects both that our natural ways of individuating belief-forming processes do not involve this kind of factivity, and also that an insistence on it will rule out some plausible instances of knowledge, most notably knowledge gained through inductive inference (Brown 2018, ch 2). On the first point, she grants that English-speakers do use the factive locution "see that" for successful cases of visual perception, and even that we have some extended use of it for the detection of mental states (e.g. seeing that another person is in pain), but she correctly observes that there is no directly parallel construction for testimony in English, and it is generally strained to apply "see that" here.<sup>5</sup> Brown notes that one might say something like "I saw she was telling me the truth" on some special occasion, but this is not our ordinary practice in reporting second-hand knowledge, and it would be outright peculiar to say "I saw that Wikipedia was telling me the truth" on having consulted that source on some minor point of trivia (Brown 2018, 33).

---

<sup>5</sup> However, it is worth noting that constructions of the form "S was told + *interrogative*" are factive; as are constructions like "S learned from X that  $p$ ." Thanks to Carlotta Pavese for this observation.

If we take our task to be one of individuating *belief*-forming processes, where it is understood from the outset that a belief is an attitude which can attach to either true or false contents, and where the aim is to predict or explain the subject's behavior or claims independently of their environment, then it is not surprising that the processes we come up with will be of types that make no essential reference to the relation between the subject and the environment. It is certainly possible to individuate belief-forming processes on the basis of the metacognitive byproducts they produce in us, and we may be particularly tempted to do this if we approach epistemological questions from the first-person perspective: from the inside, there is a conspicuous commonality between states of seeing and states of hallucinating. Still, anyone impressed by Jessica Brown's concerns about what is reflected in natural languages might worry that similar concerns also arise when we individuate belief-forming mechanisms from the inside. Do English-speakers commonly refer to a way of forming beliefs that is neutral between seeing and seeming to see? The question only becomes more pressing when we look across languages: for example, the most comprehensive survey of grammaticalized evidentials (Aikhenvald 2004) reports no language which has an evidential marking that is neutral between seeing and visually hallucinating.

However, we do not have to frame the problem from the outset as one of individuating ways of forming beliefs. Another route in to the problem would start by distinguishing ways of learning about the world; this path can make it seem more natural to focus on the distinctions between first-hand perception, second-hand learning through testimony, and sound inference. Even if similar problems arise on the side of belief-forming mechanisms individuated from the first-person perspective, it remains a good question why these factive modes of learning are not all conveniently and transparently lexicalized, across languages, especially given our reasons to believe that speakers of all languages are generally monitoring the relevant distinctions. One reason may be that our conversational contributions do not typically assert their epistemic character; even in languages where assertions must be evidentially marked, the evidential marking is not part of the at-issue content of the assertion (Simons 2007, Murray 2017).

When we do make epistemic standing explicit, some ways of learning carry more weight in common dialectical contexts than others: I can reasonably expect others to defer to me on commonplace events of which I have first-hand experience (how do you know that Bill broke the window? – I saw him do it), but it is less persuasive to respond to challenges by citing sources whose epistemic

credentials are more complex, and are equally open to the challenger (such as Wikipedia). Still, there is some evidence that we naturally think of testimony in a way which recognizes the distinction between it and pseudo-testimony: for example, note that in his initial description of the baseball game case, Baron Reed describes himself as knowing the outcome, “because my brother told me what happened”. This way of capturing the basis of the judgment is not in fact neutral between the good case and the bad: “he told me what happened” is factive, where “he told me that the Cubs won” is not. In the case where the Cubs lost, it is not true that my brother who said otherwise told me what happened. But Reed is not wrong to reach for the stronger construction in the context of a discussion of the testimonial transmission of knowledge: the tell-*wh* construction has a good claim to be our more natural way of thinking of what is happening when we speak with those we trust.<sup>6</sup> Likewise, when we use an epistemic modal like “must” to indicate inference, we are not necessarily indicating any diminished commitment to the main claim we are making, or admitting the possibility of error in our inference (Von Fintel and Gillies 2010).<sup>7</sup> In conditions where the epistemic credentials of our claims are under challenge, it may be dialectically appropriate for us to retreat to more neutral formulations concerning what we have been told, or what we gather may well be the case, but this does not mean that our initial way of thinking about the origins of our judgments is similarly cautious.

Brown’s concern about inductive inference is harder to address. It may seem odd to suggest that the progression from observing any finite subset of ravens to concluding that all ravens are black could be a process which securely locks us onto the truth, so that the truth of this conclusion is essential to the process. I will have to insist upon individuating processes so that a different process is carried out when one samples ravens in a possible world containing a few unsampled albino specimens, versus a world in which all ravens are black, and one is subject to no risk. These processes might look the same to the raven-counter, but one of them involves epistemically important luck in avoiding the white ravens, where the other was relevantly safe. In my view, the justifying basis of a judgment is not strictly a function of the choices for which one may be responsible, or even those together with the subpersonal processes involved in cognition; the basis includes environmental conditions as well. In this larger process, the evidence afforded by individual observations plays a

---

<sup>6</sup> I am grateful to Patrick Shirreff and Carlotta Pavese for discussion on this point.

<sup>7</sup> This point is controversial, as Carlotta Pavese has reminded me. For a contrary view of ‘must’, see Daniel Lassiter (2017).

crucial role in legitimizing the ultimate judgment, but it does not exhaust the normative conditions responsible for what one comes to learn through induction. Proper treatment of this issue lies beyond the scope of the present work, but I hope to have given some sense of the limited role I am assigning to consciously available evidence in my overall account of justification.

Defeat cases pose an especially interesting challenge to my account, exactly because they seem to leave the basis of a judgment fixed, while toying with the subject's representation of that basis. The next section takes a fresh look at these cases.

### **3. Defeat cases, and shifts of basis**

Roger White has a sharp statement of the problem that defeat cases pose the type of view I am defending in this chapter. White is concerned to defend a 'Cartesian' conception of evidence, in which 'the evidence I gain from visual experience consists in information about how things visually appear to me' (2014, 301). On this conception, one has precisely the same evidence in seeing or hallucinating. White contrasts his preferred view with 'Evidence Externalism', according to which 'the evidence I gain from visual experience goes beyond information about how things visually appear to me and may include facts about how things are in my environment' (2014, 301). The Evidence Externalist denies that the subject enjoys the same evidence in the Good case of seeing one's hand and the Bad Case of hallucinating a hand: the Good case includes 'a kind of direct epistemic access to the presence of a hand before me.' White then lays out a case in which a person (we will call him "Hank") who is actually looking at his hand, and is given highly persuasive but misleading evidence that he is merely hallucinating. The Cartesian view of evidence has an easy explanation of why Hank seems to lose his justification to believe that he is looking at his hand, under the pressure of the misleading evidence. The ordinarily strong relationship between the visual appearance of a hand and the presence of a hand is screened off by the apparently plausible misleading testimony about the drug. If the Evidence Externalist agrees that the misleading testimony renders Hank unable to have a justified belief that there is a hand before him, then she will need to explain how the subject's 'direct epistemic access' to outer world facts has disappeared. White argues that there is nothing available from the Externalist's depiction of the scenario to explain this shift:

First, why is it that misleading evidence that I'm hallucinating robs me of this extra evidence? Nothing appears to have changed about my perceptual state and connection to the world. Suppose I even go on dogmatically thinking *here's a hand* despite the mounting evidence that I'm hallucinating. I am at least still *seeing my hand*. My eyes are still receiving accurate information from the world and my belief forming mechanisms are operating as usual in forming a true belief. Isn't that enough to have *direct perceptual access* to the layout of my environment? I also happen to have some misleading evidence that this is not so which I am blithely ignoring. But it is hard to see how my possession of such evidence changes my *perceptual state* in such a way to rob me of my direct perceptual access to the world. (White 2014, 314-5).

White is right to observe that nothing in Hank's eyes or environment need change as the misleading reports about drugging are taken on board. What is controversial here is his claim that Hank's 'belief forming mechanisms are operating as usual in forming a true belief' as he dogmatically affirms *here is a hand* notwithstanding the apparent evidence that he is not in a good position to judge this question. By the description of the case, Hank is not engaging just the same belief-forming mechanisms on the question of his hand's presence before and after he registers the misleading evidence. In the original and naïve perceptual judgment, Hank is immediately conscious of his hand. The infallibilist account of the basis of this belief can cite such factors as the proper functioning of Hank's visual capacities and their safely true production of a noninferential judgment. To generate the impression that Hank's belief is defeated, note that it is important that he has a properly first-personal present-tense recognition of the bearing of his misleading reports about his situation on the judgment he is asked to make. Because Hank's new evidence opens the question of how he should think about whether his hand is really before him or not, he is switched from naïve to reflective cognition on that embedded question at the same time.<sup>8</sup> To understand the misleading reports, Hank must consciously entertain the possibility that he himself has ingested a drug which produces illusions as of hands; to make matters worse, he has also been given a persuasive argument that this possibility is actual. If we imagine him as 'blithely ignoring' this possibility and dogmatically judging *here's a hand*, this deliberate evasion of evidence is part of the complete belief-forming process now sustaining his belief (for another version of this argument, see Van Wietmarschen 2013, 413-15). If this process

---

<sup>8</sup> Against this kind of move, Ru Ye urges that "how one's belief is based is a purely psychological process, a process about how the belief is formed or maintained, and gaining a piece of evidence needs not change this psychological process" (2018, 7). I would urge that gaining evidence is also a psychological process, and one which forces changes in the related psychological process of answering a suitably related question. Ye maintains that defeat is ultimately best understood as a matter of intellectual responsibility; this may be right, but one could have an account of responsibility which includes the specification of responsible processes.

now involves thinking dogmatically and disregarding evidence, Hank is no longer making a judgment in a manner which generally leads safely to the truth: sound naïve thinking has been replaced by questionable reflective cognition.

Maria Aarnio offers a rival analysis of these cases. She distinguishes between the basis and the method of a judgment, where one's basis is the belief-forming process responsible for a belief, and one's method is the epistemic rule that one follows. An example illustrates the distinction:

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. (Aarnio 2010, 5)

In Aarnio's view, the stubborn subject who retains his belief in the face of misleading evidence is being unreasonable: the disposition that he is showing in insisting upon this method of belief-formation does not generally lead safely to truth, if we consider its application across a wide variety of cases. However, Aarnio argues, in this particular case the belief is still well-based and safely true: by stipulation our subject is exercising perception in favorable conditions, and his method of believing on the basis of perception is at no risk of delivering error in these conditions. The method is a good one, given the circumstances, although the subject's adoption of it shows a disposition which would in most circumstances be problematic (crime does not generally pay, but in this case the subject gets away with something). The subject's dogmatic belief actually constitutes knowledge, on Aarnio's account, and any intuitions to the contrary are to be explained by appeal to a general heuristic of taking unreasonable beliefs to fall short of knowledge, as they typically do.

Aarnio describes the stubborn subject as adopting the method of believing 'on the basis of perception.' If the basis of a belief is the whole belief-forming process responsible for its production, then arguably this moment of 'adoption' should itself constitute part of the basis of the stubborn subject's belief, a part which distinguishes this basis from that of the naïve subject. Taking a more restricted view of the basis as naïve perception, the problem is rather that this basis is no longer an option for the subject, given the stipulations of the case. It is certainly possible for the dogmatic subject to believe the same proposition that he initially believed on the basis of perception, to re-

endorse the verdict delivered to him by sight, after this verdict has been brought into question. However, if perceptual judgment is by its nature noninferential and unreflective, the stipulated presence of conscious contents that offer rival answers to the first-order question—*what is before you is perhaps not a hand, but a mere hallucination*—is enough to ensure that our subject is not judging unreflectively anymore: he now has to step through the minefield of these contents to make up his mind. Whether he likes it or not, the basis of his judgment now includes factors other than the straightforward causal sensitivity to features of the environment that constitutes ordinary visual perception. We could perhaps imagine a subject with exceptional powers of mental self-control, able to induce a selective temporary amnesia in himself to banish all the misleading evidence from his consciousness. But the agent who has somehow managed to forget all his defeaters and now looks innocently at his hand is undefeated; at the moment of judgment he is not unreasonable but naïve, and once again able to know. Accepting a proposition unreflectively is not the same psychological process as accepting it in the face of recognized defeaters.

Becoming reflective on a question ordinarily settled unreflectively does not automatically trigger epistemic self-sabotage. It is sometimes possible to wonder whether the lighting is tricky, for example, and collect more evidence to reassure oneself rationally that it is not. In such a case, one is no longer simply ‘believing on the basis of perception’; one is believing on the basis of conscious reflection which takes a succession of inputs, including perceptual inputs. As long as conscious reflection consists in sound reasoning, however, it can still support knowledgeable judgment. However, when reflection is coupled with scrutiny of the background assumptions of unreflective thought, in conditions where that scrutiny cannot be rationally addressed with the materials available to reflection, we are in trouble. One advantage of understanding defeat cases this way, rather than Aarnio’s, is that we can accept the common intuition that defeated subjects lack knowledge; we do not need to concoct an error theory.

Bob Beddor poses an interesting case, in which a subject (Consuela) sees a red vase, but is told by a “usually reliable informant” that it is a white vase illuminated by red light; however, Beddor stipulates, “Consuela (unjustifiably) regards her interlocutor as completely unreliable; hence, his testimony doesn’t causally affect her credence [in the proposition that there is a red vase before her]”. (Beddor 2014, 147-8). Beddor argues that this type of case will block efforts like mine that aim to deal with defeat through the individuation of cognitive processes. This case is somewhat more

pared-down than White's: Hank was given highly persuasive testimony, so it would be harder for him to come across as rational in discounting the odd possibility of a hallucinogenic drug out-of-hand. Consuela is simply told, falsely as it turns out, that there is a red light shining on the vase. If she understands this report, even if she does not accept it, then it is a good question whether Beddor can simply stipulate that the report has no causal impact on Consuela's credal state. To understand the report is to contemplate an explanation of one's current percepts at odds with the judgment delivered by naïve perception, at which point one would be already engaged in reflective cognition; in my view, understanding the interlocutor will already suffice for basis displacement. The interlocutor does not need Consuela to accept that the lighting is abnormal in order to raise the question of whether the lighting is abnormal, a question that cannot be settled by naïve perception, but one which bears on the legitimacy of trusting her initial impression of the colour of the vase. If Consuela closes the question about the lighting with some unjustified reasoning about the interlocutor's unreliability, then again her belief formation is compromised. Perhaps there is a way of understanding Consuela's attitude to the interlocutor as so dismissive that she does not even consider the question of whether there actually is something wrong with the lighting; on this way of reading the case, I would venture, her unreasonable attitude to this misleading interlocutor could insulate her from defeat.<sup>9</sup>

So far, I have considered cases involving a switch from unreflective to reflective cognition. Even within the realm of reflective cognition, we can generate similar problems by shifting to higher-order reflection. Reflective cognition itself involves sequentially presented contents, where the global broadcast of these contents in consciousness enables them to trigger a variety of modular subpersonal processes, including processes responsible for explicit reasoning (for a review, see Carruthers 2015). In higher-order reflection, it is possible to raise the worry that these underlying operations are going awry, so that one's ordinary capacity to reason well has been hijacked. David Christensen (2010) details a case in which one has just solved a simple logic puzzle, but then is given what seems like excellent evidence that one is under the influence of a drug which makes one generate inaccurate answers to logic puzzles of this kind, while leaving one with ordinary feelings of

---

<sup>9</sup> The insulation can only happen under rather special circumstances: perhaps the interlocutor is Consuela's little brother who teases her with suggestions she finds too irritating to consider seriously. Other ways of being dismissive do not leave the epistemic standing of her belief unscathed: if she is weirdly dogmatic about the color of the vase, for example, then Consuela's grip on the relevant first-order truths about the world is no longer safe. Thanks to Mona Simion for helpful criticism on this point.

logical competence, and an intact capacity to understand the puzzles' initial parameters. Rethinking the puzzle, Christensen observes, it seems the rational thing is to lower one's confidence that the apparently correct response is right. If the evidence about the drug is misleading, one does still have entailing grounds for one's conclusion in the well-understood initial parameters of the puzzle. However, the misled subject will be compromised in her ability to exploit those grounds. Ordinarily, rational reflection on the parameters of a logical puzzle enables the judgment that a certain answer is correct, where this answer comes to mind in a manner that seems compelling, thanks to underlying cognitive processes that execute a valid rule of inference such as *modus ponens*. These rules are executed in us by processes which are themselves unconscious, like the processes supporting syntactic judgments, although realized in different brain regions (Monti and Osherson 2012). For the subject who is conscious of the misleading evidence about the effect of the drug—an effect that produces false but apparently compelling solutions to logical puzzles—the question of the correct answer to the puzzle has been re-opened, and the fact that her initial answer to the puzzle seems compelling is now a warning sign against it. Where she could initially accept the answer on the basis of logical capacities which safely deliver the truth, where those capacities were simply exercised on the parameters of the puzzle, she is now challenged to reflectively endorse that answer while taking on board not only the parameters, but also her new apparent evidence against the hypothesis that the answer which seems compelling is right. The basis of her judgment has a new level of psychological complexity, and working in this complex manner with materials available to her, it is no longer possible for our subject to naively endorse the initial and correct answer to the puzzle, and it is no longer rational for her to reflectively endorse this answer either. Because her original reasoning was sound, the basis of her original judgment has not been rationally undermined: one could not rationally demonstrate that this flawless basis was at fault.

There is a certain elegance in understanding defeat as involving a shift in the basis of one's judgment: this strategy saves externalists from needing to add a special internalist patch to their theory of knowledge. The addition of special patches to cover defeat cases is a longstanding practice in broadly externalist analyses of knowledge, and a practice which has long been seen as problematic for these approaches. In the Classical Indian Nyāya tradition, for example, knowledge was analyzed as a 'truth-hitting cognitive episode', born of a faultless causal or justificatory source, qualified with the extra condition that the resulting judgment must be 'non-dubious in the sense that no reasonable ground for doubting its truth has appeared' (Matilal 1986, 135). Bimal Matilal notes that the 11<sup>th</sup>-

century skeptic Śrīharśa was already concerned that the ‘no-defeater’ clause seem oddly redundant, given the condition on the quality of one’s source (ibid). Contemporary Anglo-American externalists still add similar ‘no defeater’ clauses to their basic externalism, analyzing knowledge as something like reliably produced true belief in the absence of defeaters (e.g. Goldman 1986). Because the other conditions on his analysis of knowledge are often weaker, they are not always open to Śrīharśa’s worry about redundancy, but the motivations of this type of hybrid position remain difficult to explain. Laurence BonJour raises the worry that any theory of this sort will be an ‘untenable halfway house’ between internalism and externalism (BonJour and Sosa 2003, 32). When externalists consider what makes a belief normatively appropriate, the defining feature of their position is that they do not insist that the subject should always have conscious access to positive justifying grounds. It is curious, then, for externalists to add a requirement excluding the conscious availability of negative evidence. Why should conscious access matter on one side of the ledger and not the other?

If the misleading apparent evidence in defeat cases has a direct impact on our need for conscious materials, then there is an explanation of the asymmetry: by raising concerns about the assumptions of less reflective forms of judgment, defeat cases shift us upwards into more reflective ways of thinking, in which we will need consciously available materials to proceed rationally and make safely true judgments. Psychologically realistic externalists do not have to see the conscious recruitment of reasons as epiphenomenal in our belief-forming processes: operations on conscious contents are distinctive causal processes, too. Not only in defeat cases, but across the board, externalists can insist that subjects need to have appropriate consciously available input for reflective cognition, given that this type of thinking consists in sequential operations on what is consciously available. The externalist can consistently maintain that consciously available input is not needed for our unreflective forms of judgment such as perception. Maintaining psychological realism, these forms of judgment do not consist in operations on consciously available material; they simply generate consciously available results. Knowledge is generated both reflectively and unreflectively following uniform epistemic principles: no special criteria are needed to explain when we need reflectively accessible evidence, or to manage defeat cases.<sup>10</sup>

---

<sup>10</sup> For very helpful written comments on this chapter, I am grateful to Carlotta Pavese and to Mona Simion. For discussion, I am grateful to David Barnett, Maria Lasonen-Aarnio, Sergio Tenenbaum and Jonathan Weisberg. I would also like to thank the Social Sciences and Humanities Research Council of Canada for funding my research.

## References

- Aarnio, M. L. (2010). "Unreasonable knowledge." Philosophical Perspectives **24**(1): 1-21.
- Aikhenvald, A. Y. (2004). Evidentiality. Oxford, Oxford University Press.
- Beddor, B. (2014). "Process reliabilism's troubles with defeat." The Philosophical Quarterly **65**(259): 145-159.
- BonJour, L. and E. Sosa (2003). Epistemic Justification: Internalism vs. Externalism, Foundations vs. Virtues. Malden, MA, Blackwell.
- Brown, J. A. (2018). Fallibilism: evidence and knowledge, Oxford University Press.
- Carruthers, P. (2015). The Centered Mind: What the Science of Working Memory Shows us about the Nature of Human Thought. Oxford, Oxford University Press.
- Christensen, D. (2010). "Higher Order Evidence." Philosophy and Phenomenological Research **81**(1): 185-215.
- Dasti, M. and S. Phillips (2017). The Nyaya-sutra: Selections with Early Commentaries, Hackett Publishing.
- Dasti, M. and S. H. Phillips (2010). "Pramana Are Factive: A Response to Jonardon Ganeri." Philosophy East and West **60**(4): 535-540.
- Dasti, M. R. (2012). "Parasitism and disjunctivism in Nyāya epistemology." Philosophy East and West: 1-15.
- Evans, J. (2009). "Introspection, confabulation, and dual-process theory." Behavioral and Brain Sciences **32**(02): 142-143.
- Goldman, A. (1979). What is Justified Belief? Justification and Knowledge. G. S. Pappas. Dordrecht, D. Riedel: 1-23.
- Goldman, A. I. (1986). Epistemology and cognition, Harvard University Press.
- Johnson, M. K., S. Hashtroudi and D. S. Lindsay (1993). "Source Monitoring." Psychological Bulletin **114**(1): 3-28.
- Lassiter, D. (2017). Graded modality: Qualitative and quantitative perspectives, Oxford University Press.
- Matilal, B. K. (1986). Perception: an essay on classical Indian theories of knowledge. Oxford, Oxford University Press.
- Mercier, H. (2011). "On the universality of argumentative reasoning." Journal of Cognition and Culture **11**(1-2): 85-113.
- Mercier, H. and D. Sperber (2011). "Why do humans reason? Arguments for an argumentative theory." Behavioral and Brain Sciences **34**(02): 57-74.
- Monti, M. M. and D. N. Osherson (2012). "Logic, language and the brain." Brain research **1428**: 33-42.
- Murray, S. E. (2017). The semantics of evidentials, Oxford University Press.
- Nagel, J. (2015). "The social value of reasoning in epistemic justification." Episteme **12**(2): 297-308.
- Nagel, J. (2016). "Knowledge and Reliability." Goldman and His Critics. H. Kornblith and B. McLaughlin, Wiley Blackwell: 235-258.
- Papafragou, A., P. Li, Y. Choi and C. Han (2007). "Evidentiality in language and cognition." Cognition **103**(2): 253-299.
- Pomerantz, A. (1980). "Telling my side: "Limited access" as a "fishing" device." Sociological inquiry **50**(3-4): 186-198.

- Reed, B. (2012). "Fallibilism." Philosophy Compass **7**(9): 585-596.
- Schiffer, S. (2009). Evidence= knowledge: Williamson's solution to skepticism. Williamson on Knowledge. P. Greenough and D. Pritchard. Oxford, Oxford University Press: 183-202.
- Simons, M. (2007). "Observations on embedding verbs, evidentiality, and presupposition." Lingua **117**(6): 1034-1056.
- Smithies, D. (2019). The epistemic role of consciousness, Philosophy of Mind.
- Van Wietmarschen, H. (2013). "Peer disagreement, evidence, and well-groundedness." Philosophical Review **122**(3): 395-425.
- Von Fintel, K. and A. S. Gillies (2010). "Must... stay... strong!" Natural Language Semantics **18**(4): 351-383.
- Williamson, T. (2000). Knowledge and its Limits. New York, Oxford University Press.
- Williamson, T. (2014). Knowledge first. Contemporary Debates in Epistemology, 2nd Edition. M. Steup, J. Turri and E. Sosa. New York, Wiley: 1-9.
- Ye, R. (2018). "Higher-order defeat and intellectual responsibility." Synthese.