

THE SOCIAL VALUE OF REASONING IN EPISTEMIC JUSTIFICATION

JENNIFER NAGEL

jennifer.nagel@utoronto.ca

Social epistemology has to be admired for its courage in tackling those areas of human judgment that seem most epistemically problematic. When and how are we justified in accepting the testimony of a stranger? That looks like a hard question. Social epistemology has less obvious application to what might be considered the easy cases of epistemic justification: for example, the justification of judgments founded on explicit reasoning ('there is no largest prime number'), or inner sense ('I am presently feeling cold and a bit nervous'). My aim in what follows is to explore the social dimension of these 'easy' cases, and in fact also to discuss some hidden reasons why the epistemic justification of these judgments seems less problematic than the justification of judgments based on, for example, testimony and perception.

My project here is to bring together two bodies of research I consider relevant to social epistemology: work from linguistics and cognitive psychology on the structural differences between the epistemologically 'hard' and 'easy' types of judgment, and work in evolutionary psychology on the social function served by explicit reasoning. Section 1 describes the four major types of judgment distinguished in linguistic research on grammaticalized evidentials, and examines the underlying mechanism that naturally keeps track of which of these types of judgment we are making. It's an interesting question why we keep track of the judgment's type rather than simply its content; the section 2 looks for some answers to that question in a new empirical theory about reasoning, Hugo Mercier and Dan Sperber's Argumentative Theory. The third section takes a fresh look at the spectrum that runs from apparently easy to apparently harder cases of epistemic justification, in the light of Mercier and Sperber's theory. If Mercier and Sperber are right about the natural instincts driving human reasoning, we will be naturally tempted to think of epistemic justification in a somewhat unstable way. The concluding section aims to use these results to explain a corresponding instability in common philosophical approaches to epistemic justification.

I. BASIC TYPES OF JUDGMENT

Judgments are made in a great variety of ways. We could distinguish judgments made on the basis of vision, hearing, learning from controlled experimentation, mathematical proof, introspection, courtroom testimony or casual bus-stop conversations. However, some contrasts between ways of judging cut deeper than others. In what follows, I take the most basic structural divisions to lie between the following four types of judgment:

1. Inner sense
2. Perception

3. Inference
4. Testimony

Fine-grained distinctions within these categories can of course still matter to epistemology. For example, in lumping vision together with hearing and the other senses in the basic category ‘perception’, I do not mean to suggest that there is no epistemological value in exploring the differences between judgments made on the basis of different sensory modalities. Similarly, there can be epistemologically very interesting contrasts between low-stakes casual conversation and high-stakes courtroom testimony. The four identified types are basic in the sense that there is a structurally significant difference – to be explained shortly – between all perceptual judgments distinguishing them from all judgments of inner sense, inference or testimony, and so on for each of the four basic types. Even before an explicit theoretical discussion of the character of this distinction, however, you may already sense some intuitive epistemic difference between judging that the coffee beans are in the freezer because you saw them there, or because you heard someone say so, because you have inferred that this is the only place that they could now be.

The natural capacity to sense an intuitive contrast here arises from a psychological function known as source monitoring, a function which accompanies virtually all adult judgment, and manifests itself in a number of ways, including in the presence of grammaticalized evidentials in many natural languages. The four-fold division above is in fact drawn from Peggy Speas’s work on grammaticalized evidentials.¹ Roughly a quarter of the world’s languages have grammatically obligatory marking of epistemic grounding for all declarative sentences (Aikhenvald 2004). Known as evidentials, these markings have a variety of grammatical forms, including affixes, special verb forms, clitics and particles. In a language such as Tibetan, for example, in reporting the location of the coffee beans, speakers would have to use different verb endings for perceptual, testimonial, and inferential information sources; a different marker still would be used for statements based on inner sense, such as ‘I am sick’ when this judgment is made on the basis of an inner sentiment of nausea as opposed to for example reading a doctor’s report (Garrett 2001). We can of course use language to mark grounding in English as well – ‘I hear the coffee beans are in the freezer’; ‘The coffee beans must be in the freezer’; ‘The coffee beans are in the freezer, I see them’ – but it is open to us to report the location of the coffee beans while leaving the grounding of our claim unspoken, just as it is open to speakers of some languages (but not English) to make declarative statements without tense markings. This is not to say that the distinctions go untracked. Although no deliberate or consciously accessible cognitive effort is required to keep tabs on grounding, the assignment of one type of grounding or another is a natural psychological accompaniment of all normal adult judgment, in a process known as source monitoring (Johnson *et al.* 1993). Adults

1 While I follow Speas’s division of categories, I change her terminology to make it more consistent with standard practice in epistemology. So for example Speas uses ‘internal sensation/experience’ as her label for category (1), which covers our grasp of states like dizziness, nausea, moods or inner seemings; she uses ‘witnessing’ as her label for category (2), but it is clear from her examples that this category covers all and only ordinary sense perception. There is no uniform terminology in the evidentials literature; Aikhenvald (2004) uses ‘experienced’ for (1) and ‘observed’ for (2). Speas labels testimonial category as ‘hearsay/quote’ and the inferential category as ‘inference/assumption’; Aikhenvald sometimes also uses ‘reported’ for the testimonial category.

don't naturally find themselves judging that the coffee beans are in the freezer without having any idea why (we can find ourselves randomly contemplating that proposition, of course, but not judging it to be so). We can sometimes be mistaken about the grounding of a judgment, but we do naturally represent it as grounded in one way or another. For example, the knowledge that there was a giant tsunami in Indonesia on December 26, 2004 will ordinarily be represented differently for those who witnessed that tidal wave personally and for those who were not there but heard about it through the news media. Interestingly, this ability seems to develop at an equal pace in those who speak languages with and without grammaticalized evidentials (Papafragou *et al.* 2007). Researchers studying the development of source monitoring abilities in languages with and without mandatory evidential markings have concluded that 'human beings are cognitively predisposed to monitor the origins of information they come to possess', whether or not this monitoring must be expressed in language (Papafragou *et al.* 2007: 293–4).

In claiming that evidentials divide grounding into four basic types, Speas is not claiming that all languages with evidentials mark exactly those four categories, but rather that the fault lines between these four are the fundamental places where many unrelated languages are drawing lines. Many languages draw a line just between the first and last pairs: for example, Korean and Turkish have evidential markers for direct and indirect grounding, where the direct category includes both inner sense and perception, and indirect includes inference and testimony. Within the four basic types, a few languages draw further subdivisions: a number of languages, such as Tucano and Eastern Pomo, distinguish vision from other sensory modalities (Aikhenvald 2004: 51–2). Several languages, including Wintu and Tsafiki, distinguish within inference between inferences from immediately perceivable signs and from common background knowledge (Lee 1938; Aikhenvald 2004: 54). Other languages, such as Northern Embera, distinguish within testimony between a quotative evidential marking someone's exact words and a broader 'hearsay' evidential (Aikhenvald 2004: 58). However, these finer breaks generally seem to occur only within languages that are marking the larger divisions around them.² There are also firm limits to the distinctions that can be drawn: for example, there is no living language that has been clearly shown to grammatically distinguish more than five types of judgment (Aikhenvald 2004).

The cross-linguistic commonalities in evidential systems are striking: of all the various ways of making judgments that might be culturally salient, Speas notes that only a select few are ever grammatically mandatory: 'no language has an Evidential for divine revelation, experience reported by loved one, legal edict, parental advice, heartfelt intuition (gut feeling), learned through trial and error, or teachings of prominent elder/authority, for example' (Speas 2008: 944). Papafragou and colleagues observe that 'the fact that evidential systems cross-linguistically converge on the same narrow set of distinctions points to the conceptual basicness of these distinctions' (2007: 293). But why exactly are the major lines so often drawn in the same places?

A clue that Speas follows is the parallel between the grammatical marking of grounding and the marking of grammatical person. Although person marking differs among

2 There are some possible exceptions: for example, Aikhenvald reports a nearly-extinct language in Oklahoma in which there is only an auditory evidential; however, it is unclear whether this odd pattern is a function of the obsolescence of the language, or even whether the suffix interpreted as an auditory evidential has been properly understood (Aikhenvald 2004: 37).

languages, for example on whether and how gender is differentiated, all languages draw similar lines separating speaker and addressee, and between what lies within and outside the deictic sphere (the ‘here’) of the conversation. A similar geometry applies to evidential markings: they indicate the ‘location’ of the grounding of the judgment.

The parallel Speas draws between person marking and evidential markers can be illustrated with a chart showing the interaction of the two binary divisions running under each:

The basic structure of person marking (Harley and Ritter 2002):

| Person | Speaker | Deictic Sphere |
|-------------|---------|----------------|
| I | + | - |
| YOU | - | + |
| WE | + | + |
| HE/SHE/THEY | - | - |

The basic structure of evidential marking as identifying evidence location (from Speas 2004):

| Evidence | Speaker | Deictic Sphere |
|-------------|---------|----------------|
| Inner sense | + | - |
| Perception | - | + |
| Inference | + | + |
| Testimony | - | - |

A judgment grounded in inner sense has a grounding within the speaker: if I report that I am sick on the basis of my feeling of nausea, you as addressee have no immediate access to that grounding. Judgments grounded in perception are made on the basis of evidence available in the outer world present to the speaker and to others within an appropriate range: I see the spots on my arms, and so could suitably located others. Judgments made on the basis of testimony are evidentially anchored in a third-party position outside the speaker and outside what is ordinarily presumed to be immediately available to those present (‘I’m sick, the doctor says’). The trickiest geometry applies to inferential judgments, the kind of judgment formed when we say ‘Apparently John was at the party’, on surveying the telltale signs of his presence. Speas contends that these judgments are grounded in a combination of evidence located in the outer realm of what is present (we see the broken bottles on the lawn) and inner witnessing of a process of inference: while conscious of the outer cues, the speaker is also conscious of her internal process of drawing a conclusion from what she sees.

Finding a parallel between evidentials and person marking is only the first step in trying to figure out the natural purpose of our source monitoring capacity. It’s easy to see why animals have evolved to know where the coffee beans are. It’s harder to see why humans should have an obligatory system tracking how such judgments are grounded. The capacity is uniquely human: although certain non-human animals may have some ability to monitor their level of certainty, and to keep track of some of the circumstances in which they witnessed something, no non-human animal has demonstrated an ability to separate judgments formed by personal experience from judgments formed in other ways, for example by attending to the signaling of a conspecific (Carruthers 2008; Smith 2009; Crystal *et al.* 2013). Even in humans, the capacity for source monitoring is not immediately available at birth. Source monitoring is difficult enough that it takes a number of years to develop: young children are notoriously inaccurate at judging whether

their beliefs are founded on personal experience or on testimony, for example (Wimmer *et al.* 1988). Before the onset of mature source monitoring, however, the young child's awareness that the candy is in the box serves as an effective guide to action even if he is not aware whether he is judging that the candy is there because he saw it or because he was told.

What natural purpose is served by our mature source monitoring ability? One might suppose that it exists to enhance the accuracy of individual judgment. Because source monitoring gives us feedback on which cognitive process is supporting a given judgment, we can learn over time which processes are more or less reliable, as we get feedback on which judgments were mistaken coupled with how we made them. If source monitoring has an accuracy-oriented function, it could help us favor superior channels when we receive conflicting information. I think it is possible that some accuracy-related function is in fact served by mature source monitoring; however, it is unlikely that individual accuracy is an adequate explanation for our possession of this capacity. Even without consciously available source monitoring, we already have elegant mechanisms integrating varied information channels in a way that assigns them appropriate weighting for accuracy (Ernst and Banks 2002). Furthermore the divisions between inner sense, perception, inference and testimony do not map out a simple hierarchy in reliability or certainty: for example, a judgment that one is sick based on the testimony of an expert doctor could be more reliable than a judgment to the same effect based on inner sense. Similarly, the 'must be' that in English has an evidential function to mark inference does not necessarily indicate lower certainty than a report based on direct perception (Von Fintel and Gillies 2010).

Even if it can help in some circumstances with individual accuracy, the core function of source monitoring is more likely to be social, given the parallel between person marking and the evidential categories that mark the main lines of what is distinguished in source monitoring. Source monitoring matters when we need to communicate our judgments to others: indeed, even to decide what does and does not need to be conveyed, it matters where our judgments are coming from, and where our evidence is situated, relative to ourselves and our audience. Some argue for the stronger conclusion that the core function of all metacognitive operations, including monitoring of certainty as well as source, is social and communicative (e.g. Frith 2012); I'll remain neutral on whether that stronger thesis is true, and also on the question of whether source monitoring also serves some accuracy-related functions. It's enough for my purposes here that one major function served by source monitoring is social. The main aim in what follows is to look at the social value of source monitoring as a way of making sense of the epistemic division between our four basic types of judgment. To this end, the next section examines an aspect of human communication that has a direct connection to this division.

2. THE ARGUMENTATIVE THEORY OF REASONING

'Reasoning' is used to mean many things, both within philosophy and within psychology. Hugo Mercier and Dan Sperber distinguish 'inference' and 'reasoning' as follows (2011: 57):

- **Inference:** the generation of new mental representations on the basis of old ones, for example in learning from experience, in a manner that need not be conscious.
- **Reasoning:** a special kind of inference in which one moves from consciously entertained premises to some consciously entertained conclusion.

Many animals are capable of inference, in Mercier and Sperber's sense of the word. One could for example say 'after several training sessions, the pigeons inferred that rewards are given when they press the green bar after seeing the red light'. Inference in this broad sense requires no conscious or verbal thought. Reasoning, on the other hand, is something uniquely human; Mercier and Sperber point out that there are no signs of such a process in nonhuman animals or prelinguistic infants (2011: 57). It is also naturally human, in the sense that it occurs in all known human societies (for a survey of the evidence on this point, see Mercier, 2011). There is no known culture in which language is just used to describe, command or question, and never used to present explicit reasons favoring a conclusion.

The question Mercier and Sperber's theory sets out to answer is this: Why do human beings naturally engage in reasoning? Various answers have been proposed over the years, most commonly, that reasoning enhances the scope and accuracy of individual judgment. There are some difficulties with that suggestion, however. Reasoning is not needed to construct impressively accurate individual models of the world. As Mercier and Sperber point out, work comparing the performance of humans and mice in tracking complex patterns of reward, for example, shows that both species can adapt to subtly changing patterns of risk in a 'nearly optimal' manner (Balci *et al.* 2009). And while explicit reasoning does in some circumstances increase accuracy – for example in helping us avoid the conjunction fallacy in our probabilistic thinking (Tversky 1973) – in other circumstances it diminishes accuracy – for example weakening our ability to solve non-causal base rate problems (Stanovich and West 1998). The payoff in accuracy is not as clear as it should be if increased accuracy were the core function of reasoning. The accuracy explanation also fails to explain a number of natural biases in reasoning, for example confirmation bias, in which we disproportionately seek supporting reasons to rationalize pre-existing judgments, rather than instinctively employing reason neutrally to seek evidence on both sides of a question.

What Mercier and Sperber propose is that the evolutionary explanation of our capacity to reason is as follows: 'The main function of reasoning is argumentative: reasoning has evolved and persisted because it makes human communication more effective and advantageous' (2011: 60). Other animals have greatly restricted ranges of information they can communicate. Bees can signal nectar source direction, distance and strength through a very restricted set of innately programmed movements; animals like peacocks can signal their fitness to potential mates by maintaining a magnificent tail not supportable by an unfit creature. But because language gives us the possibility of expressing such a massively greater array of propositions, we face the 'cheap signaling problem': it is possible for us to express all kinds of true or false propositions very easily and at no cost to ourselves as signalers. Of course, if we can send any signal at all, whether or not it is true, then the recipients of our messages might be better off ignoring what we say, and without their uptake our signaling has no value. To be viable and persist in our species, communication must be structured in a way that will benefit both transmitter and receiver. The practice of explicit reasoning is adaptive, according to Mercier and Sperber, because it is at least a

rough solution to this problem: it enables us to share information that would not be accepted on trust. Reasoning naturally consists in the intuitively driven construction and evaluation of arguments from premises that should be accepted on trust by one's audience to conclusions that might not be. Reasoning therefore greatly extends the scope of what we can convey to each other, to the mutual benefit of members of our species.

In support of this thesis, Mercier and Sperber argue that their theory does a good job of explaining natural biases in reasoning such as the confirmation bias; they also review extensive evidence showing that argumentative contexts trigger our natural disposition to produce and follow arguments. In contexts where one needs to persuade others of something, reasoning performance is sharply enhanced, where direct incentives for accuracy typically fail to improve reasoning. Individual human performance on deductive reasoning tasks such as the Wason selection task is notoriously weak, with roughly 10% of participants answering correctly, and little improvement under financial incentives for accuracy. In small group settings, however, performance rises sharply, with roughly 80% of groups finding the solution; strikingly, very large improvements are seen even in groups where no one initially had it right (see papers surveyed in Mercier and Sperber 2011: section 2.3). Recognition of valid inferential patterns such as *modus tollens* is also sharply enhanced in argumentative settings (section 2.1). We naturally engage in reasoning in situations in which we want to persuade others of a conclusion that they would not accept on trust, but we can also be driven towards explicit reasoning even when alone, simply by the pressure of expected or imagined accountability to others (Lerner and Tetlock 1999). The argumentative theory of reasoning also does something to explain the social function of source monitoring: if we keep track of where our judgments are grounded, relative to the information sources available to our audience, then we are better positioned to use explicit reasoning when it is most needed. It is not clear that we could recognize which of our judgments would be likely to be taken on trust without some capacity for source monitoring.

Many illustrations of the function of evidentials involve contexts of persuasion. For example, Aikhenvald's discussion of the Nganasan language (spoken in Siberia) focuses on a sentence translated as follows: 'A fox must have been going round by these abandoned settlements ... it must have broken a tooth – a (broken-off) piece of tooth is lying on the ground' (2004: 48). In the original, evidentials indicating inference mark the propositions about the fox and the breaking of the tooth (but not the last observation about the tooth on the ground). Because evidentiality is lexical and optional in English, the same claims about the world could have been reported in English without the evidential use of 'must': 'A fox has been going round ... it broke a tooth ...'. The explicit use of lexical evidentials in the English translation is however appropriate as a faithful rendering of the evidentially marked Nganasan because the drawing of an inference of the sort depicted in the passage is just the kind of context in which explicit lexical evidentials are ordinarily used in English. The evidential cues help the audience sort out what is being offered as a reason for what. But whether or not the audience is to receive explicit cues of evidentiality in a setting such as this, it will matter to the hunter who seeks to communicate to be able to tell where his judgments are coming from. From the hunter's perspective, there is an important difference between conveying to your fellow hunters that you have seen the animal breaking its tooth and conveying that you infer from what you can see that the animal must have broken a tooth, and so on. Audience members can be moved to accept a conclusion about the animal's whereabouts by having their

attention drawn to visible signals; for the hunter keen to communicate the near presence of the animal it is useful that he is aware of the grounding of his own judgment that the animal is near. Non-verbal creatures could certainly respond appropriately to the traces of their prey, but for the deliberate and verbally coordinated planning of human hunters, there are clear advantages in source monitoring and the persuasive explicit argumentation it supports.

Maintaining that the natural function of reasoning is argumentative does not amount to claiming that this is its only function. Once we have the capacity to reason, however our species acquired it, we can deploy it to a great variety of ends, and we can achieve types of knowledge (for example, knowledge of abstract mathematical structures) not otherwise accessible to us. It would be a mistake to link the epistemic value of reasoning too tightly to its biological function. However, focusing on the natural basis of reasoning can help us to identify contexts in which we will have a natural instinct driving us towards the explicit production of reasons, and it can help us to keep in perspective the limitations on reasoning's capacity to provide knowledge. If reasoning is naturally activated when one considers a claim that is not likely to be taken on trust by one's audience, the contemplation of claims with a disputed basis should naturally trigger the drive to produce explicit reasons. But remembering that mice and men can know about reward patterns (and many other things) despite their inability to reason explicitly, epistemologists should keep firmly in mind that explicit reasoning is not the sole source of our knowledge. Reasoning is one potential source of knowledge, but it is also an automatic defense mechanism activated by challenges to our efforts to communicate any judgment. Mercier and Sperber's theory predicts that when what I have learned on the basis of testimony is challenged, for example, I should feel an instinct to adduce explicit reasons for its acceptance, ideally reasons that my audience will accept on trust. But my feeling this instinct is not a sign that my own initial acceptance of the claim was itself founded on explicit inference: testimonially grounded judgment is typically founded on taking a speaker's word at face value, sometimes with implicit or intuitive evaluation of trustworthiness, rather than any explicit contemplation of the speaker's trustworthiness or any explicit reasoning on my part (Sperber *et al.* 2010).

3. THE SPECTRUM FROM EASY CASES TO HARD ONES

Almost any judgment one makes may encounter resistance from an audience, but if Mercier and Sperber's theory is correct, there will be differences in the resistance that can be expected for judgments of the four different types identified in section 1. Judgments that are themselves founded on explicit reasoning from intuitively evident premises will not naturally awaken resistance, even in situations where perhaps they should. We find something instinctively satisfying in hearing the explicit production of reasons for a conclusion, even if these reasons are for example one-sided reasons supporting a conclusion we are antecedently committed to (as on our favorite cable news channels), rather than neutral surveys of the evidence for and against a given conclusion.

It is possible, if difficult, to make oneself wonder about the epistemic status of arguments from premises one finds intuitive, either by deliberately suspending one's trust in those premises, or by focusing self-consciously on the quality of one's reasoning. Although one's premises are consciously entertained in argumentative reasoning, the

structure linking those premises to the conclusion is ordinarily supplied intuitively rather than explicitly, whether we are producing an argument or evaluating one, and we have no direct introspective access to its character (Evans 2009). If explicit reasoning appears to be epistemologically ‘easy’ – if it looks like an epistemically unproblematic open book – this is in large measure because we are instinctively driven to accept the premise-conclusion patterns that others are instinctively driven to produce. As far as accuracy is concerned, judgments founded on explicit reasoning are not necessarily on firmer ground than those based on testimony or perception; on reflection, we can appreciate that explicit reasoning remains vulnerable to distorting factors such as the confirmation bias, and to the standing risk that we have taken faulty premises on trust.

Judgments of inner sense also have a sheltered social status. There is a tradition in epistemology that advocates a special epistemic status for these judgments on account of their accuracy; extreme forms of this tradition argue that judgments of inner sense are structured in a way that guarantees their correctness. The strong claims of guaranteed accuracy have been under fire recently, for example from Timothy Williamson, who argues that there is no non-trivial mental state such that being in that state guarantees that you will be in a position to know that you are (Williamson 2000). Williamson focuses in particular on the principled difficulties surrounding borderline cases of being cold, nervous or in pain. Claims of perfectly guaranteed accuracy are hard to sustain. Even granting the milder point that non-borderline judgments of inner sense are generally highly accurate, it does not seem that accuracy considerations alone suffice to motivate the special status accorded to judgments of inner sense in epistemology. But if other categories of judgment such as perception seem riskier than inner sense, this may be in part because errors in perceptual judgments are more easily noticed (by others, or by oneself at a later time), so these claims are more open to social challenge. The subject matter of judgments of inner sense largely shields them from the corrective scrutiny given to judgments about the outer world: the subject is in a privileged position (even if not an invulnerable one) relative to others with respect to the content of these judgments, and their content is of a restricted type (current internal conditions of the subject) that is ordinarily going to matter more to the subject than to the audience. If speakers can typically get their judgments of inner sense accepted on trust, these judgments will not naturally awaken the instinct to provide explicit supporting reasons.

Judgments of perception, on the other hand, make claims about a public realm accessible to others. When these judgments are challenged individually (‘Is that really a fox tooth?’) we can draw on premises of various types (testimonial, perceptual) to defend them. As epistemologists, we are also in a position to challenge these judgments collectively: how do we ever know that our judgments of perception are informative of an outer world? Our instinctive reasoning mechanisms drive us towards finding premises that must be accepted on trust in order to defend our perceptual claims. The Mercier and Sperber theory does something to explain attractiveness of the epistemological dream of finding a way to vindicate perceptual judgment in terms of a combination of inner sense and explicit reasoning. The attraction is not that arguments whose premises concern inner seemings are particularly guaranteed to be accurate, or even that they show any real promise of supporting our claims about the external world; the attraction is that such arguments promise to move from premises with a special social status in ways we generally find instinctively satisfying.

According to the empirical literature on source monitoring, we instinctively recognize some basic structural differences in the ways in which our judgments are epistemically grounded. Judgments based on perception (we see and feel the rain pouring down around us) register as distinct from judgments based on inference (we see the wet clothing and umbrellas of those entering from outside and gather that it must be raining). Our natural recognition of these differences does not give a privileged position to explicit inference in forming our picture of the world, nor does explicit reasoning have a special status in virtue of superior accuracy. However, under challenge (or the prospect of challenge) from others, explicit inference gains value. When we are concerned with persuading others, rather than simply concerned with how things are in the world, reasoning becomes crucial. Our natural attitude to justification shifts when we expect resistance from others.

This shift in our attitudes to justification lines up with a tension in epistemological internalism. The core idea of internalism, according to Robert Audi, is that ‘what justifies a belief is somehow available to the subject – through consciousness or reflection – to use in justifying it’ (Audi 2001: 22). Audi of course recognizes the distinction between being justified in believing something and having engaged in the activity of justifying it: he does not hold that prior explicit justifications are always required in order to be justified. However, he places sharp restrictions on the type of materials that will count as ‘available’ to the subject as justifiers: these materials must be ‘internal’. He elsewhere explains justification as follows: ‘justified beliefs are those one is in some sense in the right in holding: holding them is normatively appropriate given one’s sensory impressions, rational intuitions and other internal materials. In the language of desert, one does not deserve criticism (from the point of view of the effort to reach truth and avoid falsehood) for holding them.’ (29) There is an awkward relationship between the restriction of justifiers to the internal and the value placed on the effort to reach truth and avoid falsehood. It is not obvious that internal justifiers such as sensory impressions (understood as materials for judgments of inner sense) are the best materials we have to help us reach truth and avoid falsehood. What we learn from perception (understood as a function which informs us of the outer world) and testimony is indispensable in making accurate judgments about the world.

A challenge to both social and individual epistemology is to explain why we commonly think of perceptually and testimonially supported judgments as justified despite feeling worried, on reflection, that only what is internally available can justify. I’ve tried to argue here that this instability in our instincts about justification arises from a hidden slide between the demands of making judgments about the world, on the one hand, and persuading others of the truth of these judgments, on the other. Different epistemic capacities are called to the fore by each of these tasks. One natural objection to this line of thinking is that we can feel this instability in our instincts about justification even without a live audience: reflecting in solitude, it can seem to me that I don’t have immediate perceptual knowledge of the printed page in front of me, but only have an interpreted awareness of inner seemings that are compatible with the absence of the objects I take myself to see. Even if it’s true that when I speak with others, I have to resort to socially privileged materials in order to persuade them, one might wonder why similar effects emerge when I reflect on my own.

It’s not easy to determine the extent to which social effects shape private reflection, and I don’t have a worked-out view on this point. It’s possible that private reflection, especially private reflection as an epistemologist, naturally involves the anticipation of an audience:

one is rehearsing arguments, after all. It may also be worth investigating the extent to which our grasp of our own propositional attitudes is like our grasp of the attitudes of others: it has been argued that we assess our own states of knowledge and belief by turning inwards exactly the mindreading capacities we ordinarily exercise on others (Carruthers 2011). If this is true, then we could expect that raising inner challenges to the grounding of our beliefs would have effects quite similar to the effects generated when we challenge the claims of others. Deeper investigation of this issue could help to shed further light on the relationship between justification understood as a quality of our own individual judgments and justification understood as something that we can do for others when our judgments are challenged.

REFERENCES

- Aikhenvald, A. Y. 2004. *Evidentiality*. Oxford: Oxford University Press.
- Audi, R. 2001. 'An Internalist Theory of Normative Grounds.' *Philosophical Topics*, 29: 19–46.
- Balci, F., Freestone, D., and Gallistel, C. R. 2009. 'Risk Assessment in Man and Mouse.' *Proceedings of the National Academy of Sciences USA*, 106(7): 2459–63.
- Carruthers, P. 2008. 'Metacognition in Animals: A Skeptical Look.' *Mind & Language*, 23(1): 58–89.
- 2011. *The Opacity of Mind: An Integrative Theory of Self-Knowledge*. New York: Oxford University Press.
- Crystal, J. D., Alford, W. T., Zhou, W., and Hohmann, A. G. 2013. 'Source Memory in the Rat.' *Current Biology*, 23: 387–91.
- Ernst, M. O. and Banks, M. S. 2002. 'Humans Integrate Visual and Haptic Information in a Statistically Optimal Fashion.' *Nature*, 415(6870): 429–33.
- Evans, J. 2009. 'Introspection, Confabulation, and Dual-process Theory.' *Behavioral and Brain Sciences*, 32(2): 142–3.
- Frith, C. D. 2012. 'The Role of Metacognition in Human Social Interactions.' *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1599): 2213–23.
- Garrett, E. J. 2001. *Evidentiality and Assertion in Tibetan*. PhD, University of California, Los Angeles.
- Harley, H. and Ritter, E. 2002. 'Person and Number in Pronouns: A Feature-geometric Analysis.' *Language*, 78(3): 482–526.
- Johnson, M. K., Hashtroudi, S., and Lindsay, D. S. 1993. 'Source Monitoring.' *Psychological Bulletin*, 114(1): 3–28.
- Lee, D. 1938. 'Conceptual Implications of an Indian Language.' *Philosophy of Science*, 5(1): 89–102.
- Lerner, J. S. and Tetlock, P. E. 1999. 'Accounting for the Effects of Accountability.' *Psychological Bulletin*, 125(2): 255–75.
- Mercier, H. 2011. 'On the Universality of Argumentative Reasoning.' *Journal of Cognition and Culture*, 11(1–2): 85–113.
- Mercier, H. and Sperber, D. 2011. 'Why do Humans Reason? Arguments for an Argumentative Theory.' *Behavioral and Brain Sciences*, 34(2): 57–74.
- Papafragou, A., Li, P., Choi, Y., and Han, C. 2007. 'Evidentiality in Language and Cognition.' *Cognition*, 103(2): 253–99.
- Smith, J. D. 2009. 'The Study of Animal Metacognition.' *Trends in Cognitive Sciences*, 13(9): 389–96.
- Speas, P. 2004. 'Evidential Paradigms, World Variables and Person Agreement Features.' *Italian Journal of Linguistics*, 16(4): 253–80.
- 2008. 'On the Syntax and Semantics of Evidentials.' *Language and Linguistics Compass*, 2(5): 940–65.
- Sperber, D., Clement, F., Heintz, C., Mascaro, O., Mercier, H., Origg, G., and Wilson, D. 2010. 'Epistemic Vigilance.' *Mind and Language*, 25(4): 359–93.

- Stanovich, K. E. and West, R. F. 1998. 'Individual Differences in Rational Thought.' *Journal of Experimental Psychology: General*, 127(2): 161.
- Tversky, A. 1973. 'Availability: A Heuristic for Judging Frequency and Probability.' *Cognitive Psychology*, 4: 207-32.
- Von Fintel, K. and Gillies, A. S. 2010. 'Must ... stay ... strong!' *Natural Language Semantics*, 18(4): 351-83.
- Williamson, T. 2000. *Knowledge and its Limits*. New York: Oxford University Press.
- Wimmer, H., Hogrefe, G. and Perner, J. 1988. 'Children's Understanding of Informational Access as Source of Knowledge.' *Child Development*, 59(2): 386-96.

JENNIFER NAGEL is Associate Chair Graduate and Associate Professor of Philosophy at the University of Toronto. Her research focuses on knowledge and knowledge attribution, but she also has interests in metacognition more broadly, and in seventeenth-century philosophy.
