

## Mindreading in Gettier Cases and Skeptical Pressure Cases

Jennifer Nagel, University of Toronto

1. Human beings naturally ascribe states of knowledge, belief and desire to one another. From a scattered range of superficial indications—which way people are facing, what they reach for—we infer deeper states that enable us to see human motions as intentional actions, as conduct that makes sense, and can to some extent be explained and predicted, in mental terms. The inferences here are not typically a matter of conscious calculation: our imperfect but serviceable everyday mindreading is intuitive in character. For example, in recognizing a person as knowing something about his environment, we do not ordinarily engage in any deliberate enterprise of comparing the features of his situation with the requirements of some explicit theory of knowledge. In this way, ascribing knowledge is like recognizing a friend's face or seeing that she looks sad, and unlike recognizing that someone is a resident alien for tax purposes or seeing that the sum of 468 and 314 is 782. One can have the sense that someone knows or fails to know something without understanding exactly what explains one's judgment.

Intuitive judgments about the presence or absence of knowledge in particular cases do not require the prior possession of an explicit theory of knowledge. Where intuitive judgments about knowledge exhibit regular and meaningful patterns, however, it is open to us to interpret them as the products of an implicit theory of knowledge; using particular judgments as data points, we can then explicitly articulate the content of this theory. This interpretive project is delicate. Because the workings of intuitive judgment are concealed from us, we can encounter patterns of judgment that are surprising. The discovery of surprising intuitions presents us with a choice: we can construe them as either successes or failures of the implicit theory governing intuitive knowledge ascription. Successes should tell us something new about knowledge; failures should be explained away as errors of one kind or another.

Where they occur in a haphazard or unsystematic fashion, awkward intuitions can be dismissed as mere performance failures on a par with random observational errors. More systematic failures of intuition can sometimes be generated by limitations extrinsic to the cognitive system dedicated to mental state attribution. For example, just as our restricted working memory prevents vision from tracking more than five moving objects at the same time (Pylyshyn & Storm, 1988), it also limits our capacity for higher-order mental state attribution. In higher-order attribution, attributed mental states represent other mental states, as in “Alice thinks that Ben knows that Charles wants to leave”; neurotypical adults seem to have a natural limit of about five such embeddings before their mental state tracking breaks down to levels of performance no better than chance (Kinderman, Dunbar, & Bentall, 1998; Stiller & Dunbar, 2007). The discovery of restrictions of this kind poses no special problem for the reliability of epistemic intuition as contrasted with vision or other sources of evidence, and does not necessarily reveal anything particularly remarkable about our capacity to track mental states.

There is a type of failure that is for our purposes deeper and more interesting: problematic intuitions can be generated by the intrinsic limitations of our intuitive capacity to recognize states of knowledge and belief. Just as clashes between various subsystems within the visual system can generate illusions such as the Müller-Lyer effect, internal wrinkles in our mental state ascription system can give rise to paradoxical patterns of epistemic intuition. That some illusory epistemic intuitions should arise is to be expected, given the complexity of mindreading as a cognitive task and the number of subsystems involved in executing it. In sorting knowledge from mere belief we draw on a great variety of resources for calculating the evidential position of others and for decoding the significance of their behavior. For example, we seem have very specialized cognitive modules for tracking direction of gaze and shared attention to objects (Baron-Cohen, 1994). Research in developmental psychology has shown that it takes considerable time for these resources to get coordinated into mature capacities for mental state ascription. Young children do

not distinguish between knowledge and belief in exactly the same way as adults do; in particular, they have greater difficulty suppressing their own privileged knowledge of the world when evaluating the perspectives of others, and greater difficulty grasping the particular value and limits of various different ways of knowing, such as touch, sight and inference (Sodian & Wimmer, 1987; Wimmer, Hogrefe, & Perner, 1988; Wimmer & Perner, 1983). For example, if young children touch two visually indistinguishable objects with different tactile properties (like dry and wet sponges, or warm and cold cups of water), they tend to predict that others who have only seen these objects will thereby know about the tangible differences between them (O'Neill, Astington, & Flavell, 1992). Adults no longer make that particular mistake, but this is not to say that adult capacities for mental state ascription are perfectly and unproblematically integrated. For example, adults will continue to experience some difficulty in suppressing their own points of view when evaluating the beliefs of those who are more naïve, producing the bias of hindsight when evaluating one's own past state of mind, and more generally, the bias of epistemic egocentrism (Birch & Bloom, 2004; Birch & Bloom, 2007; Camerer, Loewenstein, & Weber, 1989; Mitchell, Robinson, Isaacs, & Nye, 1996; Nickerson, 1999).

It is not always easy to decide whether a strange pattern of intuition represents an intrinsic error in the intuitive system, or whether the relevant intuitive system is tracking something stranger than we had thought. When the characteristic observations of the Müller-Lyer effect clashed with other visual assessments of magnitude, our pre-theoretical grasp of the nature of magnitude was good enough to allow us to dismiss the Müller-Lyer observations as illusory even before a deeper understanding of the nature of vision gave us some insight into the reasons why we experience this effect. On the other hand, apparently contradictory visual observations of motion and rest by moving and stationary observers were not generally considered illusory, but their reconciliation required advances in our understanding of motion not achieved until the Early

Modern period, when it became evident that certain visually salient differences between rest and constant motion should not be taken at face value.

This chapter discusses two curious patterns of epistemic intuition, one on either side of the line between useful surprise and illusion. On the good side, I examine Gettier intuitions, which epistemologists are keen to embrace as telling us something new and significant about knowledge. Gettier intuitions came as a surprise to mainstream epistemologists committed to the classical analysis of knowledge as justified true belief (JTB); at least some core Gettier cases are now widely taken to be effective counterexamples to that analysis. On the side of illusion, I examine the intuitive lure of skepticism, which is widely acknowledged but rarely considered a good thing. Many epistemologists who find skeptical judgments naturally tempting in various particular cases want to resist those intuitive verdicts in favor of a theory which makes more positive room for knowledge, not least because many positive knowledge ascriptions themselves tend to register as intuitive. Oddly enough, there are some deep similarities in the underlying causes of Gettier intuitions and skeptical intuitions. These similarities make it harder than one might expect to dismiss skeptical intuitions as illusory while embracing Gettier intuitions as veridical.

Of course, it is philosophically controversial to label Gettier case intuitions as successes and skeptical intuitions as failures. Defenders of a JTB or TB analysis of knowledge might consider Gettier case intuitions to be illusory (e.g. Sartwell, 1991; cf. Weatherston, 2003). On the other side, the increased stringency triggered by skeptical pressure cases might be seen as a good thing, either by skeptics who are inclined to think that most positive knowledge attributions are mistaken (e.g. Unger, 1971) or by contextualists who take these intuitions to reflect natural shifts in the semantic content of 'know(s)' rather than mistakes on the part of the ascriber (e.g. DeRose, 1992, 1995; Stine, 1976). In what follows, I aim to defend the mainstream acceptance of Gettier intuitions and

rejection of skeptical pressure intuitions by explaining something about their natural origins.<sup>1</sup> My hope is that the project of sorting out good and bad epistemic intuitions can be aided by a better understanding of the natural mechanisms generating our impressions about knowledge and belief, just as the project of sorting out trustworthy and illusory perceptual impressions can be aided by a better understanding of the natural mechanisms involved in perception.

2. Consider the following three cases:

(A) Wanda is out for a weekend afternoon walk. She lives in a large new condominium tower downtown, and her suite is fairly small and does not have any windows that open, so she really likes to get out for some fresh air. Passing near the train station, Wanda wonders what time it is. She glances up at the clock on the train station wall and sees that it says 4:15 pm. It is in fact 4:15 pm at that moment.

(B) Wanda is out for a weekend afternoon walk near the train station and wonders what time it is. She glances up at the clock on the train station wall and sees that it says 4:15 pm. It is in fact 4:15 pm at that moment. The station clock is in fact working, but it has no second hand, and Wanda only looks at it for a moment, so she would not be able to tell if the clock were stopped.

(C) Wanda is out for a weekend afternoon walk. As she passes near the train station, she wonders what time it is. She glances up at the clock on the train station wall and sees that it says 4:15 pm. What she doesn't realize is that this clock is broken and has been showing 4:15 pm for the last two days. But by sheer coincidence, it is in fact 4:15 pm just at the moment when she glances at the clock.

The stories have slightly different content (in part because they are matched for length at 81 words each), but they paint very similar pictures of how things look to the protagonist, who in each case

---

<sup>1</sup> My account of skeptical pressure and Gettier intuitions aims to undercut some of the motivation for epistemological skepticism and for contextualism; it is not intended as a free-standing refutation of those views. A more complete attack on skepticism or contextualism would need to marshal further resources, for example, arguments in philosophy of language against the viability of a contextualist semantics (as in Stanley, 2005).

seems to be making the true judgment that it is 4:15 pm. Case (A) is the simplest: Wanda wonders about the time, and looks at a clock which is stipulated to be showing the correct time. Her judgment that it is 4:15 pm quite easily registers as a case of knowledge: in one recent study, 86% of participants who passed a comprehension screen classified the judgment this way (Nagel, San Juan, & Mar, in prep).<sup>2</sup> Case (B) is a skeptical pressure variant of case (A): the story raises the possibility that the clock is stopped, underscoring the fact that nothing in the protagonist's evidence specifically excludes that possibility, while at the same time stipulating that the clock is in fact working. Only 42% of participants who got this version of the story rated the protagonist as knowing (this was a between-subjects design, so each group of participants saw only one of the variants for each story, although all groups judged a series of different types of story). A very similar minority of participants (44%) ascribed knowledge when presented with variant (C), a Gettier case (adapted from Russell, 1948; via Scheffler, 1965).

It is not a peculiarity of this story about Wanda that the knowledge ascription ratings for skeptical pressure and Gettier versions of the story are similar: roughly comparable average ascription ratings emerged over a series of eight core stories with a variety of subject matter, where each participant encountered one variant of each story (plus a number of filler stories about different types of justified and unjustified belief formation). Overall, the eight skeptical pressure cases tested earned average knowledge ascription ratings of 39.8% (that is, 39.8% of participants who passed the comprehension screen attributed knowledge to the subject of the case, or 33.9% if one outlier case is removed). The eight Gettier variants earned average knowledge ascription

---

<sup>2</sup> The study involved 222 undergraduate participants at a suburban university campus in Canada. Each participant saw only one version of the Wanda story, among a series of similar stories and filler questions; all participants judged several stories of each variant type. Some further discussion of the study is given in (Nagel, forthcoming).

ratings of 32.9%.<sup>3</sup> These rates of knowledge ascription are markedly lower than the rates for the eight length- and content-matched simple stories of evidence-based judgment, where a majority (averaging at 72.0% across the eight stories) of the participants saw the protagonist of the ordinary case as knowing, rather than merely believing, the key proposition. For all eight core stories tested, the ordinary (A) version of the story had a higher rate of knowledge ascription than both the skeptical pressure and the Gettier version.

It is striking that participants had such similar attitudes to Gettier cases and skeptical pressure cases. If epistemologists want to embrace Gettier intuitions while distancing themselves from skeptical intuitions, it does not seem that they can do so on the grounds that the former intuitions are felt to a much wider extent than the latter.<sup>4</sup> It might seem surprising that Gettier cases do not elicit much stronger negative responses than their skeptical pressure counterparts: in story (B) the possibility that the clock is broken is mentioned but stipulated to be non-actual, and in story (C) the clock is actually broken. In this particular case the skeptical pressure variant provoked slightly stronger negative reactions than the Gettier variant (with 42 vs. 44% ascribing knowledge), but across the eight different narratives tested there was an even split on whether the Gettier or the skeptical pressure variant looked further from being a case of knowledge: for four of the eight core stories the skeptical pressure variant outscored the Gettier variant in knowledge attribution; for the other half the Gettier variant outscored the skeptical pressure variant.

What is it about cases (B) and (C) above that would make the majority of people see them as involving a failure to know? The problem is not that a momentary look at a clock cannot be a way to know the time: case (A) also stipulated that Wanda just “glances up at the clock”, and 86% of participants who got that version of the story took her to know that it was 4:15 pm. The differences

---

<sup>3</sup> It should be emphasized that it is the average, across-the-board knowledge ascription rates that are very close; there were particular cases where the skeptical pressure and the Gettier variant generated quite different knowledge ascription rates, for reasons that are not yet understood.

<sup>4</sup> Thanks to Brian Weatherson for drawing this point to my attention several years ago.

are also probably not occasioned by increased confusion for the skeptical and Gettier cases about the basic facts of the scenario: for each variant, participants were screened about what the time was and what Wanda believed about the time. Knowledge ascription data was collected only from those who passed the screens, those who agreed that Wanda had a true belief on this point. Furthermore, participants rating all three stories mostly agreed that the belief they were evaluating was justified. The average justification ratings for stories (A), (B), and (C) were 1.58, 2.32 and 1.81, respectively, on a 1-7 scale with 1.0 as “completely justified”, 2.0 as “justified”, 3.0 as “somewhat justified”, 4.0 as “neutral”, down to 7.0 as “completely unjustified”. Overall, a strong majority of participants rated the justification of protagonist’s belief positively, as at least “somewhat justified”: 94% did so for each of stories (A) and (C), and 86% for story (B).

If we are looking for a common reason why participants tended to deny that the time judgments in stories (B) and (C) amounted to knowledge, one possibility is that both stories generate a sense that the safety condition on knowledge is violated, or that although the protagonist has a true belief about the time, this belief could too easily have been wrong (Sosa, 1999; Williamson, 2000). This is one standard analysis of the stopped clock Gettier case: if Wanda had looked at the clock even a few minutes before or after 4:15 pm, she would have formed a false belief about the time. The fact that her reasonable belief is true is dangerously coincidental, or lucky in a way that seems incompatible with knowing (Merricks, 1995; Unger, 1968; Zagzebski, 1994). Meanwhile, in our intuitive response to the skeptical pressure case, something similar might be at work. Perhaps the mere mention of broken clocks elevates a reader’s impression of the level of risk involved in making a judgment like Wanda’s (Hawthorne, 2004; Williamson, 2005). This particular clock was working, but the scenario overtly reminds us that some clocks are not, and it seems Wanda could easily have been looking at some such broken clock.



Those who want to embrace Gettier intuitions while rejecting skeptical pressure intuitions could then argue that the failure of safety is real in the Gettier case, but only apparent in the skeptical pressure case. In the Gettier case (C) the clock is broken; in (B) we have only an artificially elevated sense of the risk of a broken clock. But whatever the actual risk of encountering a broken clock, it is no different for the protagonists of the basic case (A) and the skeptical pressure case (B); simply drawing our attention to the already-known fact that clocks are sometimes broken does nothing to change the fact that the clock of case (B) is working, leaving the subject of this case with the same background risk of exposure to broken clocks as the subject of plain case (A). If the judgment Wanda makes in (A) is sufficiently safe to count as knowledge, one might argue, then the judgment made by her counterpart in (B) should count that way as well. Indeed, stories (A) and (B) could describe the same person in the same objective situation,<sup>5</sup> unlike Gettier case (C), which situates its subject in an objectively different environment.

The general spirit of the safety violation theory of cases (B) and (C) may be on the right track, but its interpretation presents a number of difficulties, starting with a difficulty in understanding what is meant by 'could easily have been wrong'. One might attempt to understand that notion simply in terms of the objective chance of error. But *is* there some threshold of risk (understood as objective chance of error) above which knowledge is impossible? It is not obvious just what level of apparent risk would bar the intuitive attribution of knowledge, and it is puzzling that a series of rather different scenarios could generate a series of rather similar responses. The various stories tested in plain, Gettier and skeptical pressure variants included cases about a real deer and a lifelike deer statue; perceptual judgments of color, with the mentioned possibility or the

---

<sup>5</sup> This is not to say that precisely the same information has been presented explicitly in the two cases. Case (B) overtly stipulates that the clock is working, case (A) does not. However, in understanding case (A) we most naturally draw on the usual Gricean norms of communication in assuming that there is nothing atypical but relevant that we have not been told about the situation; we would most naturally assume that the clock is working in (A) as well.

real presence of tricky lighting nearby; facial recognition and hypothetical or actual identical twins; real diamonds and cubic zirconium fakes; social media status updates with and without deceptive intent, testimony about third parties who might or might not be different individuals with the same name, and services whose location may have changed unbeknownst to the testifier but perhaps changed back. There is a great variety of different types of possible error here, and presumably a great range of background odds of each occurring. It is not easy to see how mentioning the possibility of error could drive the apparent odds of error up just enough to bar the majority from ascribing knowledge, while permitting a majority ascription of knowledge in the plain story.

Furthermore, if we insist that the objective chance of error remains the same whether or not the possibility of error is mentioned, we might worry that consistency between our responses to these cases could equally well be achieved by shifting to a negative response to our initial case (A): perhaps the skeptic is right to deny that the ordinary person who glances at a clock ever achieves knowledge of the time by doing so. And again, focusing on the apparent objective chance of error makes it especially hard to see why responses to cases of types (B) and (C) are so similar: even if mentioning the possibility of error in (B) does inflate the apparent chance of error,<sup>6</sup> one would think the objective chance of error in (C) would still seem significantly higher, so (C) should elicit a much stronger negative response.<sup>7</sup>

**3.** Perhaps a fresh approach is called for. A review of the three stories reminds us that even on a simple question about the time, we have various more and less elaborate ways of making up our

---

<sup>6</sup> There is some reason to think it does not: there is empirical evidence (reviewed in Nagel, 2010b) that mentioning a prospect does not necessarily inflate the odds of its occurrence.

<sup>7</sup> These objections do not necessarily undermine the thought that safety is a necessary condition on knowledge: safety does not need to be understood in terms of objective chance (Sainsbury, 1997; Williamson, 2000, 2009).

minds. One might glance at a clock and come to a snap judgment. Or one might see the clock, wonder whether or not it is actually working, and take steps to check—for example, as variant (B) reminds us, one might watch the clock for a period of time. Then again, the clock might be working, but running late or early, or one might have worries about one’s eyesight: there is an open-ended list of concerns that could come to mind in the course of making a judgment about the time. Such concerns might well prompt the collection of further evidence prior to making a decision; they might also surface and then be discounted or ignored, perhaps rationally, perhaps irrationally. Our various more or less elaborate ways of thinking about a problem are known as ‘cognitive strategies’ in the literature on variations in cognitive effort (for reviews, see Broder & Newell, 2008; Payne, Bettman, & Johnson, 1993). In general, ‘low’ strategies are quick and heuristic in character; ‘high’ strategies demand greater effort and deliberate sequential consideration of various alternatives, where both the array of alternatives being considered and the methods used to adjudicate between these alternatives will naturally vary in different choice situations.<sup>8</sup> Although various more or less effortful cognitive strategies might be deployed in the course of any given judgment, limits on working memory ensure that we can’t simultaneously execute two strategies, or resolve a problem in two different ways at once.<sup>9</sup> This limitation of ours has significant consequences for our capacity to track the mental states of others.

Although the manner in which one thinks about a problem can doubtless make a difference to whether one’s answer to the problem amounts to knowledge, none of the scenarios about Wanda

---

<sup>8</sup> It is an open question whether the contrast between ‘low’ and ‘high’ is fundamentally a qualitative or quantitative one: advocates of the ‘evidence accrual’ model of variable cognition argue that ‘high’ strategies ultimately call for the same kind of thinking as their low counterparts, but set a higher threshold for evidence collection (Newell, 2005; Newell & Lee, 2010). For ease of exposition I will speak as though differences in strategy are differences in kind rather than extent, but I believe the argument of this paper could be recast in terms consistent with the evidence accrual model.

<sup>9</sup> “Resolve” should perhaps be emphasized here; on some massively modular accounts of human reasoning it might be possible for several strategies to run in parallel, with the result of the strongest winning the final competition for the attention of conscious judgment (e.g. Carruthers, 2006). There is some evidence that analytic processing can be engaged to some degree even when a contrary heuristic response is given (De Neys, Cromheeke, & Osman, 2011).

explains just how she was reasoning. The stories about Wanda all provide information about the evidence she has at her disposal, but remain silent on the issue of her cognitive strategy, or how she is thinking about that evidence. This pattern of silence on the subject's pattern of reasoning was maintained across all the stories tested by Nagel et al., and is typical of stories designed to elicit epistemic intuitions.<sup>10</sup> This is not surprising. Our intuitive mindreading system does not rely on access to the private thought patterns of others; in daily life it is activated by outwardly detectable features of their evidential position, features of the type explicitly stipulated in stories (A)-(C), or by some representation of the starting points of the subject's reasoning and relevant features of her environment. (These starting points are known as "inputs" in the psychological literature on mental state attribution; judgment is a function of strategy and inputs.) Of course we expect the supplier of scenarios to be a cooperative communicator with us, abiding by the usual conventions of rational communication (Grice, 1975). We trust that any relevant peculiarities in Wanda's thinking have been conveyed, and in the absence of any special information about her way of handling her evidence we assume that she is responding to it in some ordinary manner.

Two distinct questions can be asked about each of variants (A)-(C) above: (1) What way of thinking actually would be ordinary for a person in the circumstances described for Wanda? and (2) How do we naturally see Wanda's thinking in evaluating her? The first of these questions is relatively easy to answer. Each of the stories outlines a progression from Wanda's wondering

---

<sup>10</sup> Although they do not typically do so, it is of course possible for epistemological scenarios to detail a subject's private thought processes, and we are certainly capable of contemplating such scenarios and coming up with verdicts about them. For example, given a scenario in which it is stipulated that the subject is privately thinking about her evidence in some crazy manner to reach the belief that *p*, one could come to judge that this subject fails to know that *p*. For reasons that will be discussed in more detail below, I think that our responses to such cases would not be direct products of our intuitive mindreading system but would instead be produced by our explicit conceptual understanding of knowledge. Judgments of this type involve no mystery about why we have reached the verdict we have reached, and do not have the same potential as mindreading cases to show epistemologists something previously unrealized about the nature of knowledge. Some scenarios—e.g. the brain tumor case in Plantinga's (1993)—give information both about the evidential starting points of the subject's reasoning and about the etiology of her belief formation; for these cases only the first type of content—for example, the stipulation that the subject has no ordinary evidence for her conviction—would engage our intuitive mindreading system.

about the time to her glancing at the clock, which in all three stories looks the same to her. None of the stories stipulates that she is engaging in any explicit reasoning on the question of whether the clock is reliable, and in forgoing such reasoning Wanda would be perfectly normal. The default style of judgment in these circumstances is routine and automatic: if she has no particular reason to worry, Wanda would naturally go from looking at the clock to forming a belief about the time without any personal-level reflection on the basis of her judgment (Evans, 2007; Kahneman & Frederick, 2005; Stanovich, 2009).

The notion that Wanda would be thinking the same way in (A) – (C) squares with the traditional philosophical understanding of these cases: the interest of our differing responses to (A) and (B) arises in part from the observation that our intuitions can shift when we describe the same kind of thinking in different ways. The skeptic thinks that our response to the more detailed (B) case gives us reason to doubt our ordinary thinking across the board; the contextualist thinks that it gives us reason to suspect that the semantic content of ‘know(s)’ can shift when we mention possibilities of error. Equally, we are not supposed to conclude that Wanda fails to know in Gettier case (C) because she is thinking about the clock or its reliability in some deviant manner; we are supposed to intuit that she is justified in her belief formation, presumably making a judgment about the time in the normal unselfconscious way.

Following the reasoning of the last paragraph, we can recognize on reflection that Wanda’s thinking should be construed the same way in each of (A)-(C). But this recognition does not entail that we do in fact intuitively see Wanda as thinking the same way when we read each of these scenarios. To answer the harder question concerning how we see Wanda’s thinking in our intuitive assessment of whether she has knowledge, it will help to turn to our best available psychological explanations of how we intuitively evaluate the thinking of others. The two main approaches—

“theory theory” (TT) and simulation theory (ST)—say somewhat different things here, although I will argue that they agree on the conclusions that matter most for our present purposes.

According to TT, we have a conceptual understanding of mental states that we apply in particular situations to understand others (Gopnik & Meltzoff, 1997; Gopnik & Wellman, 1992). Information about the initial state of some person (the “inputs” to her thinking) activates the application of a causal theory of mental state formation and interaction (a Theory of Mind) in predicting and explaining her behavior. The application of this theory would not typically involve conscious personal-level inference. Our Theory of Mind is entrenched in us in a way that enables us to move intuitively from an appreciation of a subject’s starting point to a judgment about what she knows, thinks or desires, without explicit awareness of the principles that guide us, just as our ‘folk physics’ enables us to anticipate the trajectory of a falling object without our having introspective access to that theory’s principles. Small children have relatively crude Theories of Mind; over time, these crude theories are refined in response to counterevidence, although even mature theories of mind may be misapplied, for example when there is confusion about the inputs available to the person being evaluated. In addition, mature theories of mind may be systematically inaccurate about odd sorts of cases, just as our ‘folk physics’ is systematically inaccurate about certain rarely encountered patterns of motion, such as the path of an object exiting a spiral tube (McCloskey & Kohl, 1983; Saxe, 2005).<sup>11</sup>

According to ST on the other hand, it is not necessary to posit the possession of theoretical knowledge of mental state formation and interaction. Because our own minds are functionally

---

<sup>11</sup> One might expect professionally trained epistemologists to be immune from systematic intuitive inaccuracy in mental state ascription, but this will not be the case if the relevant intuitive system is cognitively impenetrable in the way naïve physics is. Even professionally trained physicists suffer the common systematic inaccuracies in intuitively plotting of the trajectories of moving objects (Kozhevnikov & Hegarty, 2001).

similar to the minds of others, we can anticipate and explain their behavior by replicating it in our own minds, rather than having to possess a theory about it (Goldman, 2006; Gordon, 1986).

Advocates of ST agree that the mindreading process begins with an appreciation of a subject's inputs; they argue that we use our own minds to simulate or model the subject's thinking, reach an judgment that is introspectively accessible to us, and project the result of our simulation back onto the subject observed. ST is particularly well positioned to explain errors of egocentrism: where our own thinking is skewed in some direction, this tendency can affect the running of our model and naturally bias our explanations and predictions concerning the thinking of others.

There are some tough questions about what exactly distinguishes representing mental states from replicating them; some quite natural ways of construing ST threaten to collapse it into being a form of TT (Perner & Brandl, 2009). There are in addition a number of hybrid views incorporating elements of simulation and theory (e.g. Nichols & Stich, 2003), and it has also been argued that something venturing beyond the existing ST and TT frameworks will ultimately be needed to capture the nature of mindreading (Apperly, 2008, 2011). However, there are certain core areas of agreement. Crucially, existing approaches agree that our mindreading system works on the assumption that given the same inputs, people will tend to think in the same way<sup>12</sup>: while variation in ways of thinking can make a difference to what people do or say, this variation will not be tracked intuitively unless it is made evident through the pattern of inputs that the subject is seen or said to collect. Although it is possible for distinct individuals to take the same inputs and arrive at different conclusions because of hidden individual differences in their ways of thinking, this type of variation will not be anticipated by our one-size-fits-all intuitive mindreading system. Indeed, the fact that a person's cognitive strategy is intuitively represented or replicated just on the basis of

---

<sup>12</sup> For evidence that mindreading is a cross-culturally universal capacity, see (Sabbagh, Xu, Carlson, Moses, & Lee, 2006; Wellman, Cross, & Watson, 2001).

an appreciation of her inputs is what explains how epistemological scenarios can generate intuitive verdicts about knowledge while remaining silent on their subjects' patterns of inner reasoning.

One might wonder how much our intuitive mindreading is compromised by its blindness to private variations in cognitive strategy. On at least one major theory of the reasons for our switches between low- and high-effort cognitive strategies, this blindness will not lower the accuracy of our everyday mindreading very substantially at all. Hugo Mercier and Daniel Sperber have argued that contexts of argument are the main condition under which we ascend to more deliberate and systematic reasoning (Mercier & Sperber, 2009; Mercier & Sperber, 2011). If we typically engage in more elaborate reasoning just when we want either to persuade others on some point, or to assess the worth of some effort at persuasion, and if our mindreading efforts in argumentative contexts are generally directed at our interlocutors, then because all parties to an argumentative context will be pushed into a more elaborate way of thinking together, we will typically be representing others accurately when we assume that they are thinking like us. Whether or not Mercier and Sperber are right about the activation conditions for systematic reasoning, if they are at least right that the kind of worries mentioned in skeptical pressure cases tend to be shared socially (“Is that clock actually working?”), then our tendency to model the thinking of others on our own will not undermine the ecological accuracy of our intuitive mindreading.<sup>13</sup>

The input-driven character of our natural mindreading generates a vulnerability, however: in those (perhaps ecologically rare) cases in which one's own way of thinking about a problem really is more elaborate than that of the subject one is observing, one's intuitive mindreading of this more naïve subject may be compromised. By contrast, we can quite easily track the mental states of

---

<sup>13</sup> The latter more modest suggestion has been supported from a variety of quarters. Boaz Keysar and colleagues, for example, point out that what is salient to one thinker will typically also be salient to others around her, and that the “burden of perspective-taking” can be socially shared insofar as differences in concerns will often be pointed out by one's audience members (Keysar, Lin, & Barr, 2003).



others who simply have different inputs but are thinking in the same way we are. For example, it is easy for older children and adults to pass standard false belief tests, such as the unexpected transfer task, in which we watch a subject seeing an object placed in a container, and continue to watch as this object gets transferred to another container while the observed subject's back is turned. In this type of task there is no difference in kind between the way the observed subject thinks about the location of the object when she sees it being placed in the first container and the way we think about it when we see it being shifted. What separates us from the subject with the false belief here is simply that we have more evidence of the object's location; we have inputs the subject does not, but they are the same kind of inputs, and we are doing the same kind of task that she is when we make a judgment about where the object is after the transfer. Consequently, our solid capacity to track inputs ensures that we are good at calculating her perspective and judging where she would believe the object is located, given what she has seen.

Unlike ordinary false belief tasks, skeptical pressure and Gettier scenarios activate the representation or adoption of a more complex way of thinking about the problem the observed subject is facing. Although the subject would actually be thinking about the problem in a relatively simple manner, we as ascribers are thinking in a more elaborate way; as a consequence, we might naturally see the subject as engaging in (but failing at) our more complicated kind of reasoning. In earlier work I described skeptical pressure cases as involving such an intuitive misrepresentation of the subject's way of thinking (Nagel, 2010b). In the skeptical pressure variant of the Wanda story, for example, discussion of the possibility of a stopped clock activates a more complex strategy for making time judgments, a sequential strategy in which a person contemplates the reliability of the clock and only subsequently makes a judgment about the time. If we see Wanda as following this higher strategy, the set of inputs she collects is intuitively inadequate for knowledge: as long as someone is actively contemplating the possibility that a clock is stopped, a momentary glance at a clock will not yield the sort of evidence she should want to make up her mind about what time it is.

There is evidence that we tend to evaluate others as sharing our privileged concerns, without being aware of our doing so (e.g. Krueger & Clement, 1994), and that this bias is very difficult to eliminate (Camerer et al., 1989). If our negative verdict on Wanda is generated by an intuitive misrepresentation of her way of thinking, however, we can reject it as illusory.

The problem is that this error theory seems to apply equally well to the Gettier intuitions we want to defend. In Gettier cases we are also thinking more elaborately than the naïve subject we are evaluating, and on the original formulation of the error theory, we would be misrepresenting the Gettiered victim as thinking like us. On the bright side, the common mechanism does explain the similarity in knowledge ascription rates for cases of types (B) and (C). But the task of defending our Gettier intuitions becomes more difficult if they are taken to arise from a distorted picture of the thinking of the evaluated subject. One might accept the misrepresentation line for both types of case and then try to argue on completely independent grounds that in Gettier cases we end up getting the right answer for the wrong reasons. Alternatively, the untoward result on Gettier cases might prompt us to check whether our error theory is overly crude: perhaps there is a slightly better way of describing what we are doing in intuitive mindreading.

Can we support a different attitude to the two types of case while still maintaining that the same mechanism underpins our response to both? If we see our task in mindreading as involving separate stages of representation or simulation and then evaluation of the agent observed, we could see both skeptical and Gettier cases as involving a failure of accuracy at the first stage: we represent or simulate simple thinking as if it were more complex than it actually is. But we do not necessarily have to see ourselves as separately reproducing and then evaluating the thought of others in judging whether they know: perhaps it would be better to describe intuitive mindreading as involving the use of a certain kind of thinking just as a standard or benchmark for the performance of the observed agent. The next (and final) section of this paper ventures a rough defense of this

second way of describing intuitive mindreading. After showing in more detail how Gettier and skeptical pressure cases manage to trigger more elaborate ways of thinking about the propositions being judged, I suggest that our own more elaborate way of thinking is not directly ascribed to the subject, but used as a benchmark in evaluating her. I then explain why the ascriber's way of thinking would be an appropriate benchmark for the standard Gettier case but not for the standard skeptical pressure case.

4. One feature that skeptical pressure and Gettier cases have in common is that both invite us to contemplate merely potential inputs—evidence that could have been collected but wasn't. Case B stipulates that "The station clock is in fact working, but it has no second hand, and Wanda only looks at it for a moment, so she would not be able to tell if the clock were stopped." When we contemplate the hypothetical possibility in which Wanda is able to tell whether the clock is stopped—which we must do in order to understand the last negated clause—we imagine a person looking at the clock for an extended period of time; however, we are told she does not look at the clock for long but instead "only looks at it for a moment", collecting a fraction of the evidence we have just imagined her collecting. This pattern is typical for skeptical pressure cases.

The other skeptical pressure cases tested by Nagel et al. all invited similar attention to evidence that might have been collected but wasn't. For example, a case about color judgments (adapted from Cohen, 2002) describes its protagonist as looking at a red table, and goes on to mention that "a white table under red lighting would look *exactly* the same to him, and Albert has not checked whether the lighting in this store is currently normal, or whether there might be a red spotlight shining on this particular table." Here again we imagine someone checking the lighting and being able to make a reasoned judgment about color conditional on the results, rather than making the simpler automatic judgment. Skeptical pressure cases about testimonially grounded

belief can draw attention to the possibility of collecting evidence about the testifier: one subject who has asked for directions is described as having “no special information on whether [her informant] is trustworthy or well-informed”, another who has called a hospital in search of information on his wife “did not stop to wonder how reliable the switchboard operator would be, or whether she might be referring to another person with the same name as his wife.” Again, understanding the road not taken—the hypothetical collection of surplus information—requires the representation or replication of a more elaborate cognitive strategy. What is more, in all of these cases, we contemplate a hypothetical scenario in which the more elaborate cognitive strategy would support a judgment contrary to the one actually made by the protagonist of the case.<sup>14</sup>

Meanwhile, Gettier cases also draw attention to evidence that goes uncollected by their protagonists. In the Gettier variant of the Wanda case, she “doesn’t realize ... that this clock is broken and has been showing 4:15 pm for the last two days.” The Gettiered Albert is in a store decorated by “a very creative lighting consultant who has put different brightly-coloured spotlights over almost all the furniture on display.” Like the visitor to Fake Barn County, Albert fails to collect evidence of the deception around him: “Most things in the store are lit so that they appear to be different colours, but not the one table Albert is now looking at. The coloured spotlights are set up so that shoppers don’t notice them at first.” Gettier cases involving testimony also invite worries of various sorts about the reliability of our informants, whether they are deliberately engaged in deception or confused and only accidentally right about the matter at issue.

---

<sup>14</sup> Mentioning additional evidence on its own is not effective in producing the skeptical effect; attention must be drawn to the possibility that a more elaborate strategy could have a contrary result. The following case, due to Jessica Brown, serves as an illustration: “Wanda is out for a weekend afternoon walk near the train station and wonders what time it is. She glances at the clock on the train station wall and sees that it says 4:15. It is in fact 4:15 at that moment. The station clock is working. Wanda could have checked this by watching the clock. Although she didn’t, her belief about the time was true since the clock is in fact working.” Here the redundancy of the additional evidence is underscored, and one feels no clear inclination to deny that Wanda knows what time it is. Thanks to Jessica Brown for the case, and for discussion of this point.

In order to make sense of these hypothetical inputs, the ascriber needs to adopt a more complex way of thinking about the problem faced by the subject. However it may be oversimplifying matters to say that the ascriber represents the subject as thinking the same way as the ascriber, only with a more restricted set of inputs. If the subject really were seen to be as anxious about the possibility of a broken clock as we are in reading case (B), then given her inputs she should also seem less confident about the time (or if she is seen as confident, she should seem to be unjustified). The lowered apparent confidence theory has been advanced by Alvin Goldman in his emendation of the account given in (Nagel, 2010b). Goldman has suggested that if our simulation of a subject's reasoning involves a projection onto her of our concerns (for example, about broken clocks), we will see the subject as less than fully confident that the time is 4:15 pm (Goldman, 2010).

It is not entirely clear that we do sense much lowered confidence here. In the research conducted by Nagel et al., none of the skeptical pressure scenarios explicitly stated that the subject formed a belief in the key proposition; all of them either outlined a situation in which it might be natural for someone to form a certain belief (as in the Wanda cases) or represented the subject as doing or saying something that would be consistent with the possession of this belief. For all scenarios, before being asked to evaluate the subject's justification and state of knowledge, participants were asked to judge whether the subject believed the key proposition (those who denied that the subject had the relevant belief were passed on to the next scenario; the knowledge ascription data comes just from participants who saw the subject as believing). On the lower-confidence theory one might have expected ascriptions of belief to drop sharply in skeptical pressure and Gettier cases, as compared to the control knowledge stories. We found instead only mildly lower rates of belief ascription: for the skeptical pressure stories 82.2% and for the Gettier cases 82.8% of participants judged the subjects to have the relevant belief, versus 91.8% for the control stories. Of course it is possible, and fully consistent with Goldman's hypothesis, that the

subjects in these cases appeared to believe, but with a level of confidence that seems too low for knowledge (cf. Bach, 2005); more work would need to be done to check whether this is so.

My original misrepresentation hypothesis is arguably also somewhat undercut by data on the perceived justification of the subjects in skeptical pressure and Gettier cases. If Wanda seems to be making a judgment about the time despite unresolved worries about the accuracy of the clock, she should come across as unjustified. However, across the board subjects in skeptical pressure and Gettier cases were rated as less, but not much less justified than their counterparts in the control (type-A) knowledge stories: mean justification rates in these conditions were 1.55 for the control stories, versus 1.91 and 2.03 for the Gettier and skeptical pressure cases respectively, on the 1-7 scale where 1 is “completely justified”, 2 is “justified”, 3 is “somewhat justified”, down to 7 being “completely unjustified”. The difference in justification ratings between Gettier and skeptical pressure cases was not statistically significant; the difference between these conditions as contrasted with the controls was significant, but insufficient to explain the variance in knowledge ascription rates.

Perhaps evaluations of knowledge are interestingly different from assessments of what a subject thinks, and whether she is justified. In the latter sorts of assessment we focus directly on the subject’s point of view; when we are asking about knowledge we look instead at the fit between the subject’s point of view and the subject’s environment as seen from our own perspective. In assessing knowledge, we need some standard for evaluating the subject’s thinking; on the proposal I will now examine, our own way of thinking about that broader environment—which would in normal conditions be shared with the subject—provides that standard.

Both Gettier cases and skeptical pressure cases draw attention to uncollected evidence and give us the sense that the subject should have collected this evidence prior to making a judgment. I have suggested that the discussion of uncollected evidence pushes us into a higher cognitive

strategy for the problem the subject faces: we move from making an automatic judgment to reasoning explicitly or sequentially about the reliability of our source. But we don't have to ascribe this higher strategy to the observed subject (even implicitly) in order to feel that she is falling short of knowing: if we intuitively take the appropriateness of our own cognitive strategy for granted,<sup>15</sup> then rather than representing Wanda as attempting but failing at our more complex way of thinking, we could more simply be sanctioning Wanda for her failure to adopt either our cognitive strategy or the range of evidence we now find intuitively necessary, given the strategy we have adopted. We could still say that we evaluate Wanda *as if* she were attempting our strategy and failing, but we do not need to see ourselves as misrepresenting her actual thinking in the course of doing so.

Because in both types of cases, we intuitively take our own way of thinking about the problem for granted, there is no immediate intuitive difference between skeptical pressure and Gettier cases. But a difference can emerge on reflection, when we think about the appropriateness of more and less complex ways of making up one's mind about the time in the objectively different circumstances inhabited by the subjects of cases (B) and (C). In skeptical pressure cases we can appreciate on reflection that the agent is succeeding at a cognitive task that can be performed very simply. In Gettier cases we can appreciate on reflection that the agent is failing to execute the more complex type of thinking that would be needed for knowledge in her environment.

To support this view of the difference between the two types of cases, one might consider the epistemic consequences of Wanda's collecting the mentioned but uncollected evidence in (B) and (C). In the scenario described in skeptical pressure case (B), if Wanda were to stop and watch the clock for a while, she could reassure herself that it is not broken; at this point she could indeed follow our higher cognitive strategy and accept that it was 4:15 pm at the moment of her first

---

<sup>15</sup> There is reason to think that we would do so, not least because the problem of strategy selection is itself in ordinary circumstances handled automatically. For discussion see (Nagel, 2010a).

glance, now arriving at this conclusion not simply on the basis of the glance but on some more complex reasoning incorporating a premise about the clock not being broken. But there is no obvious reason why a belief about the time founded in this more complex manner would in circumstances such as those specified in case (B) have a better claim to count as knowledge than a quick and non-inferential belief to the same effect.<sup>16</sup> Unless we have already been persuaded by the skeptic, we do not ordinarily think that people should only ever form their beliefs cautiously and inferentially. Moreover, throughout the hypothetical collection of further evidence, the reflective epistemologist can recognize that Wanda's original glance at the clock does not lose its original epistemic value.

In the Gettier case, on the other hand, if Wanda were to collect the mentioned but uncollected evidence she would be obliged to find a new basis for her belief about the time. The clock has been showing 4:15 pm for the past two days, and if Wanda were to be aware of this, her original glance at the clock would no longer serve her as any indication of the time.<sup>17</sup> In the objective circumstances in which she finds herself, a more effortful cognitive strategy is needed to secure knowledge of the time.<sup>18</sup> Epistemologists can recognize on reflection that a cautious cognitive strategy is required for knowledge in the Gettier case, not that the subject can be faulted as irresponsible for her failure to adopt such a strategy. From our superior vantage point we know that things are not as they seem, but there is nothing in her environment that should raise suspicion

---

<sup>16</sup> For further discussion of this point, see (Nagel, 2010a).

<sup>17</sup> Many Gettier cases have the defeater structure of the stopped clock example, but there are other types of Gettier case, such as Gil Harman's dead dictator case (Harman, 1973), which have some significant differences. Notably, in the dead dictator case, the subject's original basis is sound, but the subject seems to lack the ability to defend his judgment against misinformation that is (unbeknownst to him) widely circulated in his environment. These cases also involve the mention of uncollected evidence, and the threat that further evidence collection would support a contrary verdict on the key proposition. I think an extension of the present account to such cases could be developed, but I will not attempt it here.

<sup>18</sup> For some further discussion of the relationship between knowledge and strategy selection, see (Nagel, 2010a).



either about the proposition she has judged to be the case, or about the cognitive strategy she has pursued.

We do not typically engage in explicit reasoning about the level of cognitive effort we should apply in the course of making a given judgment; indeed it has been pointed out that if cognitive strategy 'selection' were itself a matter of conscious choice, we would find ourselves in a vicious regress of thinking about how hard to think (Rieskamp & Otto, 2006). We cannot always know how much effort will be needed in a given situation in order to secure knowledge, and in some circumstances we apply extra effort needlessly or fail to try as hard as we should. Sometimes greater effort is applied in response to challenges or concerns from our interlocutors; sometimes greater effort is applied when we discover some reason to think that there is some problem in the basing of a belief. Epistemic intuition does not always distinguish between needed and unnecessary strategy shifts, but on reflection we can have second thoughts about possible gaps between how hard a subject seemed to be trying and how hard he had to try.

On the account advanced here, one of the blindspots of intuitive judgment generates a class of epistemic illusions: because our mindreading system does not directly track variations in cognitive strategy, we are intuitively insensitive to these differences. When we read case (A), it is easy to make an automatic judgment about the time, and to judge Wanda's automatic judgment by the standard of the strategy we have applied ourselves. In reading cases (B) and (C), we are reasoning in a more complex manner, and judge Wanda accordingly, although we can recognize on reflection that Wanda has no particular need to reason our way in case (B).

Epistemic and perceptual illusions have something in common. In classifying some perceptual experiences as illusions, we recognize that vision does not give us perfect access to objective qualities such as spatial magnitude, but we are not plunged into skepticism about perception. We can restrict our trust in an intuitive or perceptual system and work towards

understanding the features of this system that produce the illusions we encounter. If we are prone to illusory epistemic intuitions, then knowledge itself can differ from what we intuitively register as knowledge, just as spatial magnitude can differ from what vision presents as magnitude. But this result should not be taken to suggest that we would be better off investigating the nature of knowledge without reliance on epistemic intuition. The role played by intuitive knowledge ascription in our efforts to understand the nature of knowledge may well remain as vital as the role played by sense perception in our efforts to understand the nature of space.<sup>19</sup>

---

<sup>19</sup> For detailed comments on an earlier version of this chapter, I am grateful to Jessica Brown, Mikkel Gerken and Angel Pinillos. For helpful criticism and discussion I'd also like to thank Michael Blome-Tillman, Patrick Rysiew, Valerie San Juan, Sergio Tenenbaum and audience members at the October 2010 Knowledge Ascriptions Workshop at Arché, University of St. Andrews, Scotland.

## References

- Apperly, I. (2008). Beyond Simulation-Theory and Theory-Theory: Why social cognitive neuroscience should use its own concepts to study. *Cognition*, 107(1), 266-283.
- Apperly, I. (2011). *Mindreaders: The Cognitive Basis of "Theory of Mind"*. Hove and New York: Psychology Press.
- Bach, K. (2005). The Emperor's New 'Knows'. In G. Preyer & G. Peter (Eds.), *Contextualism in Philosophy: Knowledge, Meaning, and Truth* (pp. 51-89). New York: Oxford University Press.
- Baron-Cohen, S. (1994). The eye direction detector (EDD) and the shared attention mechanism (SAM): Two cases for evolutionary psychology. In C. Moore & P. Dunham (Eds.), *The Role of Joint Attention in Development* (pp. 41-59). Mahwah, NJ: Erlbaum
- Birch, S., & Bloom, P. (2004). Understanding children's and adults' limitations in mental state reasoning. *Trends in cognitive sciences*, 8(6), 255-260.
- Birch, S., & Bloom, P. (2007). The curse of knowledge in reasoning about false beliefs. *Psychological Science*, 18(5), 382-386.
- Broder, A., & Newell, B. (2008). Challenging some common beliefs: Empirical work within the adaptive toolbox metaphor. *Judgment and Decision Making*, 3(3), 205-214.
- Camerer, C., Loewenstein, G., & Weber, M. (1989). The curse of knowledge in economic settings: An experimental analysis. *The Journal of Political Economy*, 97(5), 1232-1254.
- Carruthers, P. (2006). *The Architecture of the Mind: massive modularity and the flexibility of thought*. New York: Oxford University Press.
- Cohen, S. (2002). Basic knowledge and the problem of easy knowledge. *Philosophy and Phenomenological Research*, 65(2), 309-329.
- De Neys, W., Cromheeke, S., & Osman, M. (2011). Biased but in doubt: Conflict and decision confidence. *PloS one*, 6(1), e15954.
- DeRose, K. (1992). Contextualism and Knowledge Attributions. *Philosophy and Phenomenological Research*, 52(4), 913-929.
- DeRose, K. (1995). Solving the Skeptical Problem. *Philosophical Review*, 104(1), 1-52.
- Evans, J. (2007). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, 59, 255-278.
- Goldman, A. (2006). *Simulating minds: The philosophy, psychology, and neuroscience of mindreading*. New York: Oxford University Press.
- Goldman, A. (2010). Philosophical Naturalism and Intuitional Methodology: the Romanell Lecture 2010. *Proceedings and Addresses of the American Philosophical Association*, 84(2), 115-150.
- Gopnik, A., & Meltzoff, A. (1997). *Words, Thoughts, and Theories*. Cambridge, MA: MIT Press.
- Gopnik, A., & Wellman, H. (1992). Why the child's theory of mind really is a theory. *Mind & Language*, 7(1 2), 145-171.
- Gordon, R. (1986). Folk psychology as simulation. *Mind & Language*, 1(2), 158-171.
- Grice, H. P. (1975). Logic and conversation. *Syntax and Semantics*, 3, 41-58.
- Harman, G. (1973). *Thought*. Princeton, NJ: Princeton University Press.
- Hawthorne, J. (2004). *Knowledge and Lotteries*. New York: Oxford University Press.
- Kahneman, D., & Frederick, S. (2005). A model of heuristic judgment. *The Cambridge handbook of thinking and reasoning*, 267-293.

- Keysar, B., Lin, S., & Barr, D. (2003). Limits on theory of mind use in adults. *Cognition*, 89(1), 25-41.
- Kinderman, P., Dunbar, R., & Bentall, R. (1998). Theory-of-mind deficits and causal attributions. *British Journal of Psychology*, 89(2), 191-204.
- Kozhevnikov, M., & Hegarty, M. (2001). Impetus beliefs as default heuristics: Dissociation between explicit and implicit knowledge about motion. *Psychonomic Bulletin & Review*, 8(3), 439.
- Krueger, J., & Clement, R. (1994). The truly false consensus effect: An ineradicable and egocentric bias in social perception. *Journal of personality and social psychology*, 67(4), 596-610.
- McCloskey, M., & Kohl, D. (1983). Naive physics: The curvilinear impetus principle and its role in interactions with moving objects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9(1), 146-156.
- Mercier, H., & Sperber, D. (2009). Intuitive and reflective inferences. *J. St. BT Evans & K. Frankish (Eds.), In two minds: Dual processes and beyond*. Oxford, UK: Oxford University Press.
- Mercier, H., & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences*, 34(02), 57-74.
- Merricks, T. (1995). Warrant entails truth. *Philosophy and Phenomenological Research*, 55(4), 841-855.
- Mitchell, P., Robinson, E., Isaacs, J., & Nye, R. (1996). Contamination in reasoning about false belief: An instance of realist bias in adults but not children. *Cognition*, 59(1), 1-21.
- Nagel, J. (2010a). Epistemic Anxiety and Adaptive Invariantism. *Philosophical Perspectives*, 24, 407-435.
- Nagel, J. (2010b). Knowledge ascriptions and the psychological consequences of thinking about error. *Philosophical Quarterly*, 60(239), 286-306.
- Nagel, J. (forthcoming). Intuitions and Experiments: A defense of the case method. *Philosophy and Phenomenological Research*.
- Nagel, J., San Juan, V., & Mar, R. (in prep). Gettier Case Recognition.
- Newell, B. (2005). Re-visions of rationality? *Trends in cognitive sciences*, 9(1), 11-15.
- Newell, B., & Lee, M. (2010). The Right Tool for the Job? Evidence Accumulation in Decision Making. *Psychological review*.
- Nichols, S., & Stich, S. (2003). *Mindreading: An integrated account of pretence, self-awareness, and understanding other minds*: Oxford University Press, USA.
- Nickerson, R. S. (1999). How we know--and sometimes misjudge--what others know: Imputing one's own knowledge to others. *Psychological Bulletin*, 125(6), 737-759.
- O'Neill, D., Astington, J., & Flavell, J. (1992). Young children's understanding of the role that sensory experiences play in knowledge acquisition. *Child Development*, 63(2), 474-490.
- Payne, J., Bettman, J., & Johnson, E. (1993). *The Adaptive Decision Maker*. New York: Cambridge University Press.
- Perner, J., & Brandl, J. L. (2009). Simulation a la Goldman: pretend and collapse. *Philosophical Studies*, 144(3), 435-446.
- Plantinga, A. (1993). *Warrant and proper function*: Oxford University Press, USA.
- Pylyshyn, Z., & Storm, R. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism\*. *Spatial Vision*, 3(3), 179-197.

- Rieskamp, J., & Otto, P. (2006). SSL: A theory of how people learn to select strategies. *Journal of Experimental Psychology: General*, 135, 207-236.
- Russell, B. (1948). *Human Knowledge: its scope and its limits*. New York: Simon & Schuster.
- Sabbagh, M. A., Xu, F., Carlson, S. M., Moses, L. J., & Lee, K. (2006). The development of executive functioning and theory of mind. *Psychological Science*, 17(1), 74.
- Sainsbury, R. (1997). Easy possibilities. *Philosophy and Phenomenological Research*, 57(4), 907-919.
- Sartwell, C. (1991). Knowledge is merely true belief. *American Philosophical Quarterly*, 28(2), 157-165.
- Saxe, R. (2005). Against simulation: the argument from error. *Trends in cognitive sciences*, 9(4), 174-179.
- Scheffler, I. (1965). *Conditions of Knowledge*. Chicago: Scott, Foresman.
- Sodian, B., & Wimmer, H. (1987). Children's understanding of inference as a source of knowledge. *Child Development*, 58(2), 424-433.
- Sosa, E. (1999). How to defeat opposition to Moore. *Nous*, 33(s13), 141-153.
- Stanley, J. (2005). *Knowledge and Practical Interests*. New York: Oxford University Press.
- Stanovich, K. (2009). Distinguishing the reflective, algorithmic, and autonomous minds: Is it time for a tri-process theory. *J. St. BT Evans & K. Frankish (Eds.), In two minds: Dual processes and beyond*. Oxford, UK: Oxford University Press.
- Stiller, J., & Dunbar, R. (2007). Perspective-taking and memory capacity predict social network size. *Social Networks*, 29(1), 93-104.
- Stine, G. (1976). Skepticism, relevant alternatives, and deductive closure. *Philosophical Studies*, 29(4), 249-261.
- Unger, P. (1968). An analysis of factual knowledge. *The Journal of Philosophy*, 65(6), 157-170.
- Unger, P. (1971). A defense of skepticism. *The Philosophical Review*, 80(2), 198-219.
- Weatherson. (2003). What good are counterexamples? *Philosophical studies*, 115, 1-31.
- Wellman, H., Cross, D., & Watson, J. (2001). Meta analysis of theory of mind development: The truth about false belief. *Child Development*, 72(3), 655-684.
- Williamson, T. (2000). *Knowledge and its Limits*. New York: Oxford University Press.
- Williamson, T. (2005). Contextualism, Subject-Sensitive Invariantism and Knowledge of Knowledge. *The Philosophical Quarterly*, 55, 213-235.
- Williamson, T. (2009). Probability and Danger. *The Amherst Lecture in Philosophy*, 4, 1-35.
- Wimmer, H., Hogrefe, G., & Perner, J. (1988). Children's understanding of informational access as source of knowledge. *Child Development*, 59(2), 386-396.
- Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13(1), 103-128.
- Zagzebski, L. (1994). The inescapability of Gettier problems. *The Philosophical Quarterly*, 44(174), 65-73.