

The Mystery of Stakes and Error in Ascriber Intuitions

Wesley Buckwalter

Research in experimental epistemology has revealed a great, yet unsolved mystery: why do ordinary evaluations of knowledge-ascribing sentences involving stakes and error appear to diverge so systematically from the predictions professional epistemologists make about them? Two recent solutions to this mystery by DeRose (2011) and Pinillos (2012) argue that these differences arise due to specific problems with the designs of past experimental studies. This chapter presents two new experiments to directly test these responses. Results vindicate previous findings by suggesting that (i) the solution to the mystery is not likely to be based on the empirical features these theorists identify, and (ii) that the salience of ascriber error continues to make the difference in folk ratings of third-person knowledge-ascribing sentences.

Imagine that two spouses are waiting to deposit a check at the bank on a Friday evening and the lines inside are very long. One says to the other, “I know the bank will be open tomorrow,” and suggests they return on Saturday morning instead. This mundane knowledge *assertion* seems true. However, now suppose instead that their check was actually for a large sum of money meant to cover a series of impending overdue bills, making its immediate deposit extremely important for their financial futures. It even occurs to them that the bank might have altered the open hours since their last Saturday visit. One says to the other, “I don’t know the bank will be open tomorrow.” In the latter circumstance, this knowledge *denial* seems true. Many philosophers have claimed that these contrary intuitions are best explained by the idea

that nonepistemic factors like practical stakes or the salience of error can play an important role in everyday evaluations of knowledge-ascribing and knowledge-denying sentences.

The precise way of accounting for intuitions in these famous bank vignette pairs (see DeRose 1992) has played a prominent role in contemporary epistemology. One recent trend is to cite the content of our intuitions in bank and bank-style cases to support philosophical analyses of knowledge based on the way we actually make and assess knowledge ascriptions. Thus the ability to explain how the same knowledge denial and knowledge assertion can both be true in these two different contexts has inspired support for some of our best new competing theories of knowledge. And while debate continues among contemporary philosophers, most agree that this ability to account for bank intuitions—or the purported fact that people are more likely to ascribe knowledge when stakes and error are low than when they are high—counts as one significant theoretical advantage.

Parallel to this theoretical work in epistemology, experimental philosophers of the last few years have also begun to investigate the factors that influence ordinary language practices involving knowledge sentences (for a review, see Buckwalter 2012). They've done this by running experiments in the social psychological tradition by constructing stimulus materials closely resembling bank case vignettes. In doing so, these researchers have sought to understand the conditions and the mechanisms at work when we make third-person attributions of knowledge. However, the results these experimental philosophers have uncovered have inspired a great, and yet unsolved, mystery. Current data suggest that people's answers in these kinds of cases systematically diverge from the predictions of trained epistemologists.

With few exceptions, the evidence to date indicates that when participants in an experiment are presented with bank-case-style stimuli, the factor of stakes plays only a marginal role in their knowledge ascriptions to others.¹ This is puzzling since on the one hand, most professionals have agreed that stakes play an important role in how we assess knowledge-ascribing sentences. But on the other hand, experimentalists have been unable to reproduce commensurate findings. So, could it be that these experiments are somehow deficient in detecting the mechanisms that buttress our ordinary language practices involving knowledge, and if so, how? Or alternatively, are theoretical

intuition pumps in the hands of epistemic experts less accurate measures of everyday knowledge ascriptions, and if so, why? To answer these questions is to solve *the mystery of ascriber intuitions*.

Recently, two promising solutions to the mystery have been proposed by DeRose (2011) and Pinillos (2011, 2012) along the former lines, by questioning the designs of past experimental studies. They have suggested that folk intuitions only *appear* to radically diverge from professional epistemologists regarding subject stakes and attributor error due to various problems with the particular experiments that have been used to examine them. In response to these claims, this chapter presents two new experiments designed to directly test the solutions put forward by DeRose and Pinillos. After making adjustments for their worries, results vindicate previous findings by suggesting that (i) the solution to the mystery is likely not based on the particular empirical features these theorists have identified, and (ii) that the salience of ascriber error continues to make the difference in folk ratings of third-person knowledge-ascribing sentences.

In order to uncover the true culprits in our caper of stakes and error in ascriber intuitions, the chapter proceeds as follows. Section 1 reviews the evidence traditional philosophers and experimental epistemologists have generated involving third-person mental state attributions of knowledge. Sections 2–3 introduce the solutions proposed by DeRose and Pinillos, and respond with experiments designed to test them. Section 4 discusses possible resolutions to the mystery of ascriber intuitions based on these new data, as well as advances two further hypotheses about its origins. Lastly, Sections 5–6 conclude by revisiting the implications such results have for two leading theories of knowledge supported by ordinary intuitions.

1 Professional intuitions and experimental data

It is often taken for granted in the contextualist literature that people's intuitions in bank cases fluctuate between *low cases*, where the stakes or error possibilities of the case are minimal, and *high cases*, or where stakes and error are critical. Such intuitions have inspired some epistemologists to develop analyses of knowledge by focusing precisely on ordinary ascribing behaviors.

The result is that in addition to traditional factors like evidence or justification, recent discussions of knowledge have also begun to include more practical considerations that seem to be at work in everyday judgments.²

One strategy to capture these purported intuitions between bank cases is to focus on the pragmatic conditions that are directly relevant to the subject of the bank vignettes. Theorists such as Hawthorne (2004), Fantl and McGrath (2002, 2010), and Stanley (2005) advance accounts of knowledge by observing that not only the truth-relevant factors, but also the practical factors of the main character's situation seem to have a large effect on our bank-case intuitions. Specifically, the claim is that ascriptions of knowledge are sensitive to the practical interests of a subject, the importance that he or she is right, and the personal costs involved with being wrong. While each of these theories are subtly different, for the sake of simplicity we will refer to the common metaphysical view, *interest-relative invariantism* (IRI), as the view that roughly, whether or not a true belief counts as knowledge, depends in part on what is at stake for the *subject*.

A different strategy to capture bank intuitions is to focus on the use of the word “knows” and the factors that influence the truth of knowledge-ascribing sentences. Standard *epistemic contextualists* (DeRose 1992, 1999, 2005, 2009; Cohen 1988, 1999, 2004) claim that in everyday usage, the very same knowledge sentence often seems to be true in specific sorts of contexts but false in others. While different contextualists are free to debate the factors that influence these truth conditions, the two factors most often discussed are the stakes of the *attributor* when ascribing knowledge to a third person, and also, whether or not the possibilities for error have been made salient for the attributor from context to context. The resulting semantic thesis regarding instances of “knows that p” is that the truth conditions of knowledge-ascribing sentences can be different from one conversational context to another, based on those details of the *attributor's* situation.

One thing that interest-relative invariantists and advocates of contextualism have in common is that they both claim to capture ordinary language practices involving knowledge-ascribing sentences. To find out what these ordinary practices actually are, evidence is usually collected as follows. We are invited to consider vignettes, or pairs of fictional scenarios, in which all the features of the cases are fixed, and where we are asked to make a decision involving what

the protagonist knows. Then between the cases, we vary the factors of stakes or error, to see if our intuitions about the character's knowledge changes. When we do this, if the result is that our intuitions change from case to case, this suggests that these factors play a role in how we assess knowledge sentences. The results of such thought experiments in the philosophical literature have indicated that most philosophers agree that stakes play an important role in ordinary knowledge ascription.³

And beyond the careful manipulations presided over by today's leading epistemologists in these specific cases, there is something generally intuitive about the idea that the costs of one's beliefs weigh heavily on our assessments of doxastic states.⁴ If this is the case, then it should be a straightforward task to empirically pinpoint stakes effects in our behaviors. And indeed, there are now several results in the social psychological literature surrounding this general issue of stakes and error independently of research on knowledge specifically. Mayseless and Kruglanski (1987), for instance, show that people's subjective probability estimates are affected by the desire to avoid judgmental mistakes, in view of their perceived costliness. Fischer et al. (2008) have shown that people make decisions involving loss with significantly less subjective decision certainty (that loss-decisions are more difficult to make than decisions involving gains), and that this effect of loss-framing systematically increases people's desire to search for confirmatory information like evidence. Similarly regarding error, hypotheses of selective exposure (see Smith et al. 2008) have long since held that people tend to avoid dissonant, but seek out consonant information, suggesting that the salience of error might affect our capacity and motivation for processing evidence in the face of critical information.

However, in these discussions it is very important to distinguish the epistemic ascriber, or the person who claims someone has knowledge, from the epistemic subject, or the person of whom these claims are made. What the social psychology literature suggests is that regarding the latter, subjects in high-stakes situations, or cases of high potential for error, will have significantly less confidence in their judgments.⁵ What these results do not directly speak to is the former, or how such factors might affect third-person knowledge attributions, or our evaluations of other people's first-person knowledge assertions.

This further question of third-person attribution is exactly what experimental philosophers have attempted to investigate.⁶ Despite the intuitions of epistemologists, as well as the results from social psychology, these data in experimental philosophy have largely turned up negative results for stakes. As Schaffer and Knobe (2012) say when summarizing these findings, “This research suggests . . . that—contrary to what virtually all of the participants in the contextualism debate have supposed—neither stakes nor salience impacts the intuitions of ordinary speakers.” The majority of these findings involve the familiar bank stimuli discussed above. While each experiment is slightly different, they all involve manipulating stakes by varying the importance of a protagonist’s financial circumstances (for instance, whether or not they have an impending bill coming due) and error factor by varying whether a character proposes some kind of relevant alternative to the knowledge claim asserted (for instance, that banks can sometimes change their hours without notice).

Four independent studies have investigated the way stakes and error influence folk knowledge practices in these particular cases. Buckwalter (2010) found that in cases where a subject in this situation makes the knowledge claim, “I know the bank will be open on Saturday,” participants found it to be true no matter the stakes or error of the case. Feltz and Zarpentine (2010) found that whether or not epistemic subjects made knowledge assertions in a low-stakes cases or knowledge denials in high-stakes cases had no effect on people’s judgments about the truth or falsity of the knowledge statements. On the other hand, when participants were specifically asked to ascribe or deny rather than evaluate stated knowledge sentences, May et al. (2010) were able to detect a small effect of stakes (though mean judgments suggested that participants still agreed that subjects had knowledge despite this difference). Lastly, Schaffer and Knobe (2012) were able to detect a significant effect for error by making the error possibilities more salient through vivid and personal anecdotes (and we will revisit this idea in Section 2.2).

Of course, while all of these studies involve the same general bank context, they each have their differences (for instance, they ask slightly different questions or use distinctive stimulus materials). Overall, however, people tended to either strongly ascribe knowledge or judge knowledge statements to be true despite the stakes manipulation of these various bank experiments.

Only one study (Schaffer and Knobe 2012) was able to detect an effect for error. So at first glance, experimental data in philosophy appear to be at odds with evidence from three different sources: general intuitive seemings one might have about the role of stakes, the results contextualists and interest-relative invariantists have uncovered in select cases, and also, some neighboring results in social psychology concerning the impact of stakes on subjects' degrees of confidence.⁷ But how could this be?

2 Bank experiments and epistemic contextualism

Bank data generated from empirical studies conducted by experimental philosophers and from the intuitions of traditional epistemologists are at odds, and the game's afoot to explain why. We now turn to the first candidate solution given by DeRose (2011), according to which three important deficiencies in the experimental designs of past studies are said to be responsible.

2.1 DeRose challenges the data

In "Contextualism, Contrastivism, and X-Phi Surveys," DeRose raises several crucial objections against previous bank experiments, as well as the implications these data might have for standard epistemic contextualism. DeRose (2011) points out that the studies all fail to incorporate all of the crucial elements needed to successfully test the empirical predictions of contextualism. The result, claims DeRose, is that "it turns out the intuitive support for contextualism doesn't really face much of a wave of empirical trouble," and that "there are some severe problems with taking the results of any of these studies (whatever their aims) as undermining the intuitive basis for contextualism" (83).

According to DeRose, the first problem with earlier research is that some studies (May et al. 2010; Schaffer and Knobe 2012) use stimulus materials in which the epistemic subjects did not make an explicit knowledge statement within the confines of the actual vignette. Instead, these researchers asked participants to rate whether or not they thought subjects had knowledge in the various cases. The worry is that in cases where participants are asked to ascribe knowledge, rather than evaluate a knowledge statement, they

will be more likely to consider the particular features of their context (the experimental setting) rather than the epistemic subjects' context (as stipulated in the vignette). So, if further research were to explicitly test the predictions of contextualism, a better experiment would have participants evaluate the truth of knowledge statements made by a speaker in the actual vignette (as opposed to simply being asked whether a given character in that vignette has knowledge).

A second worry raised by DeRose is that previous experiments (Buckwalter 2010) fail to test all the crucial factors said to influence ordinary truth judgments of knowledge sentences. While it is true that some contextualists predict differences based on attributor stakes or the possibility for error, others suggest that it is the *interaction* of these two factors that affects attributors' conversational contexts. And as DeRose says in an earlier paper, "The best case pairs will differ with respect to as many of the features that plausibly affect the epistemic standards, and especially those features which most clearly appear to affect epistemic standards as possible" (2009, 54). In the name of simpler experimental designs, however, bank cases are often presented by comparing the effect of stakes or error individually. By using vignettes that isolate these two factors, the concern is that the test pairs will not have included all of the relevant factors contextualists claim influence epistemic standards. So a better test would include cases with combinations of both high stakes and salience (as opposed to just one or the other).

Lastly, DeRose points out that the best test cases for contextualism should capture how ordinary speakers actually use the knowledge claims in question. Typically, people in situations with low epistemic standards will tend to ascribe knowledge, and people in high-standards situations will deny knowledge. But some previous experiments asked participants only about the truth conditions of assertions (Buckwalter 2010; May et al. 2010). So the worry, however, is that by excluding people's evaluations of knowledge denials, we will have accidentally manipulated some property of knowledge assertion, and not stakes or error more generally. Here, one might appeal to what David Lewis called the "rule of accommodation" (Lewis 1979). According to this rule, if someone says something false in a given conversational context, we will seek to change the features of that context (via any number of specific context-altering factors) such that the new context makes the

content of that utterance true. Supposing that “knows” is a context-sensitive term subject to Lewis’ rule, it’s possible that participants in the bank experiments are more likely to agree that a *high* knowledge denial is true than they are to agree that a *high* knowledge attribution is false simply because of accommodation, and not because they sincerely judge high conversational contexts irrelevant when evaluating knowledge sentences. So in order to test this claim, participants in future experiments would need to receive cases involving both knowledge assertion and denials across the various conversational contexts presented.⁸

DeRose claims that while previous bank experiments may have included just one of these various components, only cases meeting all of these challenges will serve as accurate measures for the intuitive basis of contextualism. In the meantime, the absence of any one of these features could well serve as the explanation to the mystery of ascriber intuitions. So it looks as though experimental philosophers need to go back to the lab in order to more accurately test whether the predictions made by contextualists about stakes and salience in bank cases have ordinary intuitions on their side.

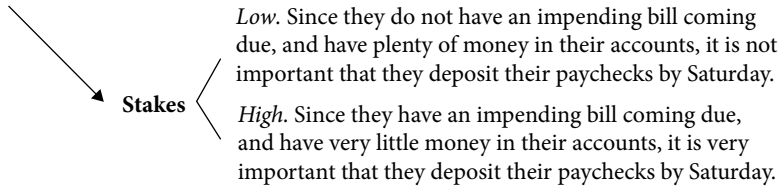
2.2 Meeting DeRose’s challenges

The worries by DeRose discussed in the last section serve as important challenges to the interpretation of previous bank results. So one straightforward strategy is to move forward by including these challenges as factors in new experimental studies. The following experiment was performed to further test contextualism in precisely this way, by importing DeRose’s three worries into a new research design point by point. This is accomplished by independently varying three critical factors within the basic bank-style vignette: (i) the *speech act* of the character in the vignette could either be one of asserting that she has knowledge or denying that she has knowledge, (ii) the *stakes* could either be high or low, and (iii) the *error possibilities* could either be salient or nonsalient. And given the findings of Schaffer and Knobe (2012), this manipulation of the salient error possibilities differed from other previous studies by using a *concrete and vivid* example of error.

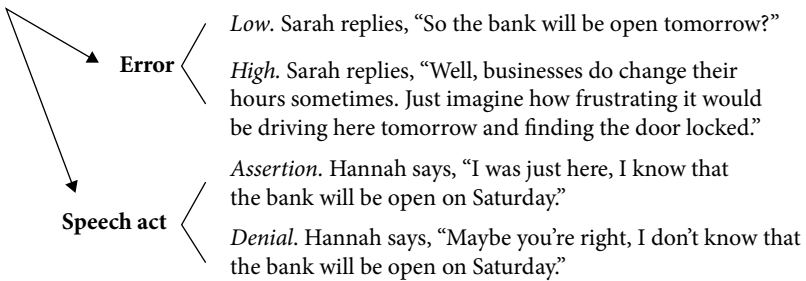
This resulted in a study with a 2 (stakes) × 2 (error) × 2 (speech act) between-subjects design in which each participant was randomly assigned to

one of eight possible conditions. The possible combinations of the resulting eight conditions are shown below.⁹

Hannah and her sister Sarah are driving home on a Friday afternoon. They plan to stop at the bank on the way home to deposit their paychecks. As they drive past the bank, they notice that the lines inside are very long, as they often are on Friday afternoons.



Hannah says, "I was just at this bank two weeks ago on a Saturday morning, and it was open till noon. Let's leave and deposit our paychecks tomorrow morning."



After seeing one of the possible bank case combinations, and receiving a pair of comprehension checks, participants ($N = 215$, 32 percent male) were then asked the following question:¹⁰

Assume that as it turns out, the bank really was open for business on Saturday. When Hannah said, "I (*know/don't know*) that the bank will be open on Saturday," is what she said true or false?

Answers were assessed on a five-item scale anchored with truth-value terms (e.g., 1 = false, 3 = in between, 5 = true). Mean truth-value judgments in the bank cases are represented in Figures 6.1 and 6.2.

The study yielded three key results.¹¹ First, there was a main effect for speech act whereby people thought that knowledge sentences involving an assertion, no matter the error or stakes of the case, were more likely true than knowledge sentences involving denial.¹² However despite this main effect, it's

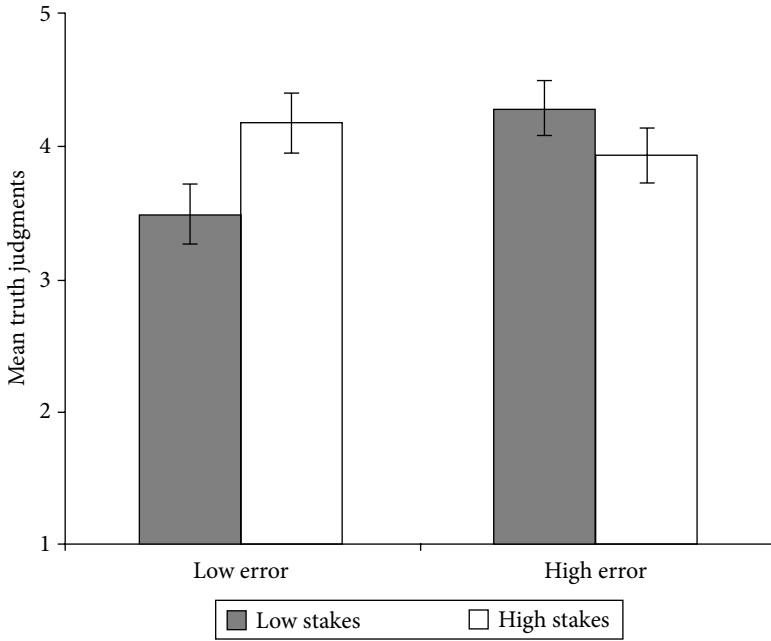


Figure 6.1 Mean truth judgments for knowledge *denials* grouped by error (\pm SE, scales ran 1–5).

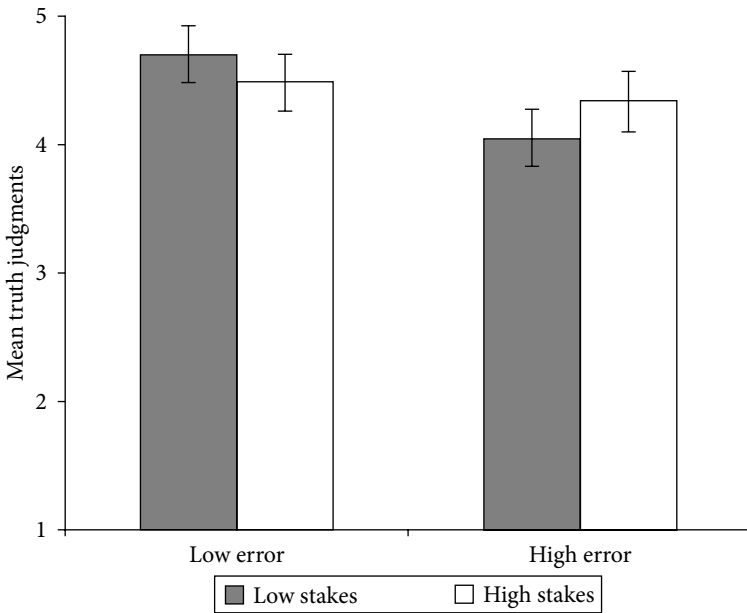


Figure 6.2 Mean truth judgments for knowledge *assertions* grouped by error (\pm SE, scales ran 1–5).

also true that participants in the experiment just generally judged everything true across the board. Notice that in the graphs above, responses have not been recoded from their original values. But since a “5” in knowledge assertion conditions is logically equivalent to a “1” in knowledge denial conditions, we can clearly see how high truth-value judgments in Figures 6.1 and 6.2 indicate that participants gave very different responses between speech act types. This suggests that while people find the knowledge assertion true in *low* and a knowledge denial true in *high* as contextualists predict in bank cases, intuitions were largely driven by accommodation.

Second, we find exactly the impact of error possibilities that standard epistemic contextualists predict. When the possibilities for error were made salient (in a concrete and vivid way), participants were more inclined to say that the assertion of knowledge was false and that the denial of knowledge was true.¹³ This effect is shown below in Figure 6.3 by collapsing across all the various levels of high- and low-stakes cases administered:

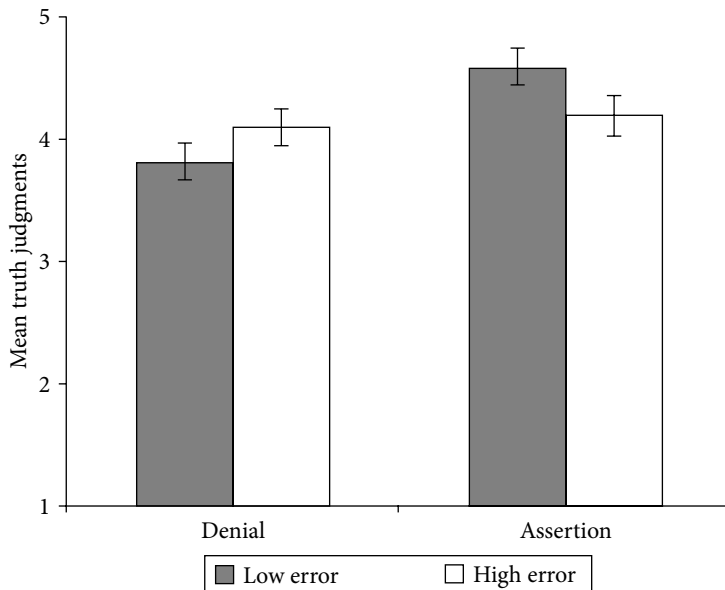


Figure 6.3 Mean truth judgments for low- and high-error conditions grouped by *assertion* and *denial* (\pm SE, scales ran 1–5).

Lastly, despite obtaining the predicted effect of attributor error, the predicted effect of subject stakes was not found.¹⁴ There was no general tendency for people in the high-stakes bank conditions to be more inclined to think that

assertions of knowledge were false or that denials of knowledge were true.¹⁵ Even after correcting for all of the previous worries in bank case stimuli, an effect for stakes was not detected in people's evaluations of these particular third-person knowledge-ascribing sentences.

So here's the state of play. Previous experiments have demonstrated that there is little reason to think stakes or error play meaningful roles in folk assessments of knowledge ascriptions in bank cases. Then, DeRose proposed three objections to the extant ascription data that might be responsible for the incongruent results between philosophers and ordinary people. When incorporating those exact worries into the current test (along with the vivid error possibilities advocated by Schaffer and Knobe 2012), we found that the actual problem with previous tests had to do with the error manipulation. That is, data seem to show that when error is vividly presented to participants in cases of both attributions and denials, there is an effect on third-person ascription.¹⁶ So the key message from this further experiment on the bank cases seems to be that the possibility of error made salient to the attributor does have an impact on the evaluation of truth conditions of sentences that attribute or deny knowledge, but that subject stakes do not.

Of course, it is often difficult to confidently interpret null stakes results. Generally speaking, there could be numerous different reasons for why a given experiment does not find a hypothesized effect. While one reason could be that no such effect actually exists, any number of experimental confounds could also be responsible. One could also object that people simply weren't paying attention, or that they were not holding fixed the important epistemic features of the cases (like the amount of evidence possessed by an epistemic subject, for instance) and that is why an effect of stakes on knowledge was not found on this particular occasion. While these are important worries to consider, it is not clear that they serve as viable explanatory hypotheses of the data in hand in the current bank study.

Importantly, the space of candidate explanations of the null results for stakes in the present experiment is constrained by the interaction effect of error possibilities and speech act type. Unlike the previous studies that have turned up negative or null results for both stakes and error, this study found the predicted effect for the latter but not the former in a single experiment across the very same stimulus materials. So a plausible explanation for the absence of

the impact of stakes would still need to retain the ability to explain this result for error. Therefore it seems unlikely that one could rely on accusations about epistemic attention or shifting evidence to explain the lack of impact stakes had on knowledge judgments, while participants simultaneously behaved exactly as epistemologists were predicting when it comes to error. In this regard, we can have even more confidence than before to think that participants are considering these cases as epistemologists intend. It's just that they don't arrive at the *same epistemic intuitions* as the experts in these particular cases regarding stakes as they do for error.

3 Evidence-seeking experiments and IRI

While there now might be some doubt about the specific features of the cases above regarding contextualism and salience of error, these data seem to very clearly question the ability of interest-relative invariantist positions to capture ordinary intuitions regarding subject stakes in bank cases. We now turn to the second candidate solution given by Pinillos (2012) suggesting that problems specifically to do with the stakes manipulations of previous bank-case experiments can account for the mystery of divergent ascriber intuitions.

3.1 Pinillos challenges the data

Pinillos (2012) offers some compelling new experimental evidence suggesting that unlike the bank data consensus in experimental philosophy, subject stakes do in fact influence third-person mental state attributions of knowledge. Specifically, Pinillos uses an experimental paradigm measuring the amount of evidence participants require an epistemic subject to collect before they ascribe knowledge to that subject. The stimulus involves a college student who is proofreading his assignment for typos. In a low-stakes condition, it is not particularly important that the assignment has no errors, while in a high-stakes condition, the student faces disastrous consequences should even one error be discovered by his professor:

Typo Low Stakes. Peter, a good college student has just finished writing a two-page paper for an English class. The paper is due tomorrow. Even

though Peter is a pretty good speller, he has a dictionary with him that he can use to check and make sure there are no typos. But very little is at stake. The teacher is just asking for a rough draft and it won't matter if there are a few typos. Nonetheless Peter would like to have no typos at all.

Typo High Stakes. John, a good college student, has just finished writing a two-page paper for an English class. The paper is due tomorrow. Even though John is a pretty good speller, he has a dictionary with him that he can use to check and make sure there are no typos. There is a lot at stake. The teacher is a stickler and guarantees that no one will get an A for the paper if it has a typo. He demands perfection. John, however, finds himself in an unusual circumstance. He needs an A for this paper to get an A in the class. And he needs an A in the class to keep his scholarship. Without the scholarship, he can't stay in school. Leaving college would be devastating for John and his family who have sacrificed a lot to help John through school. So it turns out that it is extremely important for John that there are no typos in this paper. And he is well aware of this.

After seeing one of these conditions, participants were asked, "How many times do you think [*Peter/John*] has to proofread his paper before he knows that there are no typos?". They were then told to fill in the blank with the number they thought was appropriate.

The study showed that when participants were presented with either the low-stakes or the high-stakes conditions, people thought that the student needed to collect more evidence in order to know there were no typos when the stakes of the case were high (median = 5) than when they were low (median = 2). The finding suggests that since people require more evidence before ascribing knowledge to epistemic subjects in this way, folk attributions of knowledge must be sensitive to stakes.

Furthermore, Pinillos suggests that this experiment may give us a unique perspective on what went wrong in bank cases. The worry, claims Pinillos, is that there really is no way to track whether participants are holding fixed crucial details of the cases, such as how much evidence the epistemic subject has between conditions. The thought is that evidence-seeking experiments are better equipped in this regard, since it is precisely the amount of evidence the subject should have which is measured. By detecting fluctuations in the amount of evidence required, these experiments show differences in knowledge judgments by stakes, casting doubt on the divergent intuitions previous

experimental philosophers have detected in bank cases. So regarding theories like IRI, it looks as though experimental philosophers also need to return to the lab, in order to find out exactly how stakes are influencing judgments about evidence-seeking behaviors when ascribing knowledge.

3.2 Meeting Pinillos' challenges

One latent worry in the evidence-seeking design is that the differences detected in the amount of evidence collected between low- and high-stakes subjects in this particular experiment could arise not because third-person mental state attributions of knowledge are or aren't intrinsically sensitive to stakes, but rather because high-stakes subjects are expected to collect more evidence than low-stakes subjects to actually have an outright belief on the issue at all. So a further study was run to answer the question of whether stakes specifically affect mental state ascriptions of knowledge, or alternatively, whether these differences are instead an effect for some other mental state, like belief.¹⁷

In this study, 100 participants were given a manipulation as close as possible to what is used in Pinillos' study involving subject stakes, but also varied the kind of mental state ascription that was attributed to that subject.¹⁸ This resulted in a 2×2 between-subject experimental design, independently varying the practical stakes (either high or low) and the mental state (either belief or knowledge) of the subject.¹⁹ After seeing one of the same stimuli given above in Pinillos' original experiment, participants were then asked, "How many times do you think Peter has to proofread his paper before he [*believes/ knows*] that there are no typos?" and to "Please insert the number you think is appropriate in the space below." Results are represented in Figure 6.4 below by mean scores of the amount of evidence needed in each case:

We find that as before, stakes had a huge impact on ascriber intuitions. However, while the experiment showed a significant difference on people's judgments between low- and high-stakes contexts, the specific mental state they were asked about within these contexts did not.²⁰ In other words, participants gave roughly the same answers when they were asked how much evidence was needed to be collected before the epistemic subject had knowledge as they did how much evidence needed to be collected before the subject had a belief that a certain result would obtain.

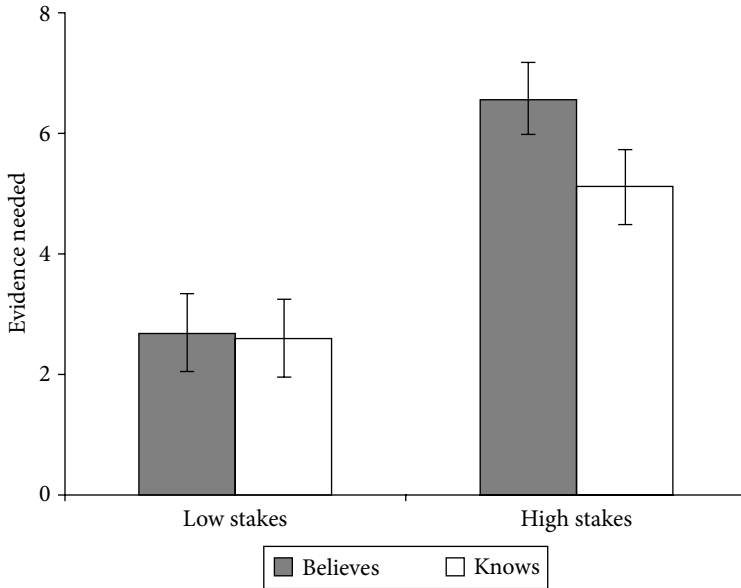


Figure 6.4 Typo case results grouped by stakes and mental state (\pm SE).

In the dispute between intellectualists and interest-relative invariantists, advocates of IRI usually hold that subject stakes are themselves supposed to bear directly on the criteria for whether a subject's true belief constitutes knowledge. Yet, identical scores between subjects in questions regarding whether an epistemic subject knows something is the case and believes something to be true suggest that the effect shown for stakes does not tell us about the specific criteria people are actually using when deciding if a subject's belief constitutes knowledge. Instead, the fact that participants found that high-stakes subjects are expected to collect more evidence than low-stakes subjects do in order to count as believing suggests that subjects in high-stakes conditions are expected to collect more evidence just to make up their minds at all. It's not that both subjects' preexisting beliefs are transformed into knowledge with a greater amount of evidence in high-stakes cases than low-stakes cases, but rather that epistemic subjects need a disproportionate amount of evidence when forming the requisite belief.

This suggests that this particular evidence-seeking experiment is a fine-grained enough measure to have detected that subject stakes have some kind of an effect on participants' general responses in the current experiment, but not

fine-grained enough to show that the relevant effect reveals something specific about knowledge in particular.²¹ For better evidence supporting the empirical predictions made by IRI, data would need to demonstrate that the stakes of the case are what matter for a subject's true belief to count as knowledge.²² To be clear, nothing in this response to Pinillos rules out the possibility that stakes do actually play such a role in people's judgments, or that such evidence could be collected in the future. Instead, this experiment was designed to show that we are not warranted in inferring support for IRI from these particular data regarding stakes. Further research is necessary before supporters of IRI can reasonably make the inference that knowledge ascription is particularly sensitive to stakes in the relevant way. And, antecedent empirical reasons for the conclusion that stakes do not play this role should inspire caution before overturning previous results to the contrary.

4 Toward solving the mystery

We began with a great mystery. How can we make sense of the systematic differences between professional intuition and folk judgments in bank cases? Data indicate it is unlikely the mystery can be entirely solved by appealing to the specific problems that DeRose and Pinillos have identified with previous experiments. However, neither does the evidence suggest that philosophers using more traditional methods were completely mistaken about actual knowledge practices. Joining with past research, further data on bank-case intuitions suggest that while error is a factor that influences ordinary third-person mental state ascriptions of knowledge, stakes are not. The latest experiments show that knowledge ascriptions in the bank cases fluctuate when vivid error possibilities are made salient. Data from evidence-seeking experiments on the matter of stakes are shown to be inconclusive, and without further testing, do not yet undermine the growing consensus in experimental philosophy that subject stakes play but marginal roles in attributor judgments.

Given these data, it seems that the solution to the mystery of ascriber intuitions is twofold. Something went wrong when previous experimental epistemologists claimed that the salience of ascriber error does not affect

people's knowledge judgments, and something went wrong when professional epistemologists claimed that their intuitions about the importance of subject stakes *actually reflect ordinary people's evaluations* of knowledge-ascribing sentences. While the empirical data are a good start to solving this mystery, an interesting further question remains why experimentalists and traditional philosophers made the mistakes that they did. Before going on to speak of the philosophical ramifications of the data in these cases, we will pause to hypothesize about why or how this mystery developed.

One hypothesis to explain why previous bank-case studies were unable to detect the intuitional variance epistemic contextualism predicts between contexts of low- and high-ascriber error seems relatively straightforward. In past experiments on bank cases, the factor of error was manipulated by only minimally mentioning the possibility of error. In Buckwalter (2010), for instance, the high-error bank-case interlocutor challenges her epistemic subject by speculating on only one general way in which she could be wrong (e.g., "Banks are typically closed on Saturday. Maybe this bank won't be open tomorrow either."). Similarly, May et al. (2010) take a similar tack when constructing a case of high error by having the epistemic subject point out that generally speaking, banks do change their hours. However, current empirical evidence suggests that the effect was detected in the present experiment simply by making the particular error manipulation in the bank vignettes more vivid: "Just imagine how frustrating it would be driving here tomorrow and finding the door locked." Following Schaffer and Knobe, who showed a salience of error effect for knowledge attribution, present results demonstrate a similar effect when participants are asked to make truth judgments of knowledge ascriptions.²³

The first culprit then, that helps explain the advent of the mystery of ascriber intuitions, is a specific problem with earlier experimental research. The problem was that the error possibilities were not made salient enough in bank cases to detect the difference in intuition between low- and high conversational contexts.²⁴ Regarding the disagreement about the factor of stakes, however, it still remains incredibly puzzling why the intuitions found in these experimental studies continue to diverge so systematically from both the intuitions of trained epistemologists, as well as the predictions that results concerning first-person confidence might have made for third-person

knowledge ascription. Further experimental evidence continues to support the hypothesis that knowledge is sensitive to error, while continuing to question the joint hypotheses that practical stakes, and the link between subjects' degrees of confidence and stakes, have anything but marginal impacts on ascribers' intuitions in bank cases.

One hypothesis to explain this difference is that ordinary people and trained epistemologists approach these thought experiments in different ways (see, e.g., Phelan forthcoming). On the one hand, participants of an experiment usually experience one particular case, and are then asked to report an immediate intuition about whether or not the epistemic subject has knowledge. By contrast, trained philosophers often proceed by considering pairs of bank vignettes together and then engage in a kind of reflection about whether the relevant differences in context have any epistemic importance. The philosophers then go on to make predictions about the judgments ordinary people will make in these cases on the basis of that evaluation. And these two different types of approaches to epistemic judgments may be shaped by two very different kinds of psychological processes. The former seems to be an implicit system-one intuition-generating capacity that enables us to respond to epistemic intuitions in particular cases with which we are presented. The latter is a system-two process, involving a more abstract set of theoretical beliefs about epistemic principles, as well as predictions about how others might conform to those principles.²⁵

So could these different decision-making approaches between these two different groups account for the mystery in bank cases? Indeed, it's possible that the predictions made by philosophers are subject to a kind of *distinction bias* often discussed in behavioral economics (Chatterjee et al. 2009; Hsee and Zhang 2004, 2010). The basic idea is that when presented with several vignettes differing by stakes—in similar to what Hsee and Zhang call “joint evaluation mode”—professional philosophers identify this feature of the vignettes and then make choices and form predictions based on their training and knowledge of the abstract epistemic principles involved. But conversely, the processes underlying ordinary people's judgments when they are presented with singular cases—or in “single evaluation mode”—are based on preferences related to actually experiencing those particular cases. And since it seems likely that the preferences that philosophers use in the former mode of evaluation

will be different from those of nonphilosophers in the latter, these different modes of evaluation may encourage philosophers to *overpredict* the impact of stakes in people's actual knowledge judgments in bank cases.

So regarding the mystery of ascriber intuitions, perhaps the second culprit is the bias that arises from the combination of formal philosophical training, together with making predictions when cases are presented under joint evaluation rather than experienced. If true, this hypothesis may be able to help explain how judgments made by expert epistemologists gave rise to the importance of stakes, but also how the role of stakes in ordinary bank-case judgments was mispredicted or exaggerated.²⁶

5 Implications and philosophical importance

Though we have made some empirical progress in resolving the mystery of ascriber intuitions, the mystery's denouement raises perhaps an even more complicated philosophical question: how does this explanation about stakes and error in bank cases bear on epistemic contextualism and IRI?

Beginning with contextualism, many philosophers have argued that as a semantic theory, it makes particular linguistic commitments about word usage. Hawthorne (2004) and Stanley (2005) argue, for instance, that the relevant kind of context-sensitivity of "knows" is objectionable because our usages of that particular word frequently deviate from usages of other common indexicals said to be context-sensitive. Going even further, Brown (2013) argues that since contextualism provides a linguistic model for "knows," and given that such models provided by leading theories of contextualism are committed to certain kinds of context-sensitivity, contrary behavioral data about folk knowledge practices regarding such sensitivities would actually threaten to undermine the view entirely.

If these arguments are correct, then getting the right results in bank-case experiments seems crucial. As it stands, one outcome of research up to this point is that semantic theories of standard epistemic contextualism are supported by experimental data showing an effect for at least one specific kind of context-sensitivity. Particularly, data show that such views that wish to include a correct theory of people's ordinary language practices regarding

sensitivity to conversational contexts should focus not on the practical stakes, but rather on *the error possibilities made salient to the attributor*. In such cases, contextualism would not be undermined by the current experimental results. The clear upshot is not only has contextualism been shown to be compatible with the relevant knowledge behaviors, but also that such empirical evidence can be used to forge more detailed versions that specify with greater accuracy the relevant linguistic model claimed for the word “know.”²⁷

Unlike the data relevant to contextualism showing folk sensitivity to error possibilities, however, current results continue to suggest that third-person attributions were insensitive to stakes. And, such findings seem to be clearly at odds with the premise that IRI best explains bank intuitions. In response to this tension, Brown (2013) argues that such experimental data only threaten to undermine one popular way of arguing for IRI, but not the position itself. Following Brown, we might note that there is no necessary dependence of the truth of metaphysical theses regarding things like temporal parts, the nature of substance—and likewise the determinants of knowledge—on concordant folk judgments or practices regarding their central theoretical entities. As a metaphysical theory of knowledge, the truth of IRI does not turn on, and is not committed to ordinary intuitions about knowledge *per se*. Therefore, despite the results of the current studies, interest-relative invariantists are free to continue to include premises about the epistemic roles of subject stakes for the metaphysical determinants of knowledge at the cost of ordinary language.²⁸

What the empirical evidence does seem to suggest, *so far* at least, is that IRI may not provide the best explanation for our epistemic behaviors regarding stakes in bank cases and beyond. In other words, the data from Pinillos have not been enough to convince us that ordinary assessments of these particular types of knowledge-ascribing sentences count in favor of the metaphysical view that knowledge is stakes-sensitive. And, while a metaphysical view need not enjoy folk agreement to be true, a safe bet is that the conclusions of such a thesis are, generally speaking, more likely true when supported by true premises rather than false ones. So this may encourage future supporters of IRI to develop and embrace alternative arguments for their view based on something other than folk practices regarding stakes (again see, e.g., Brown 2013; Fantl and McGrath 2010). Of course, there’s always the possibility that future experiments will

discover the long-lost case that does display persistent stakes effects. And such cases will have to be evaluated on their own merit as they arise. But at the very least, the current data generated in response to DeRose and Pinillos—joined with the difficulty several independent researchers have faced in detecting anything but negligible stakes effects—begin to question whether building an epistemology around folk stakes sensitivity is a very good idea.²⁹

6 Conclusion

Experiments continue to suggest that accommodation and the salience of ascriber error, but not subject stakes, makes the difference in the ordinary evaluation of third-person knowledge sentences. But research exploring the ways in which people actually evaluate knowledge sentences can still be a benefit to the more traditional research in the field, serving to help supplement, and not supplant, such methods when appropriate. One of the main goals of this chapter has been to show that the empirical investigation of the mystery of ascriber intuitions can help contextualists and interest-relative invariantists become the best versions of themselves.

By suggesting which specific features of an attributor's context ordinarily affect the standards for knowledge, this research begins to allow a more accurate estimate of the linguistic model of the word "knows." The result is that one way to *be a better contextualist* is to develop versions of the theoretical view in which the context-sensitivity of "knows" varies by accommodation and error. Similarly, experiments also continue to question the evidence supporting the claim that ordinary third-person knowledge judgments are sensitive to subject stakes. Such results may undermine one particular way of arguing for the thesis that knowledge is stakes-sensitive. Yet, they may also suggest that one way to *be a better interest-relative invariantist* might be to accept versions of the view that do not rely on premises concerning ordinary knowledge practices that people may not—or may not prevalently have. In both cases, empirical research in epistemology is additive to the theoretical work and methods in the field, inciting new directions for the partnership between future theoretical and experimental work on contextualism and IRI.

Acknowledgments

Special thanks to James Beebe, Keith DeRose, Mikkel Gerken, Josh Knobe, Josh May, Jennifer Nagel, N. Ángel Pinillos, Jonathan Schaffer, Jason Stanley, Jonathan Weinberg, and other blog members who participated in lengthy discussions on the *Certain Doubts* blog, for helpful comments and suggestions. I am grateful to Josh Knobe, Jesse Prinz, and Stephen Stich for insightful comments on previous drafts, and continued support.

Notes

- 1 Since the writing of this chapter, new work by Sripada and Stanley (2012) has claimed to detect a stakes effect in unrelated cases. However, a critical discussion of these recent findings will be saved for a later occasion (see Buckwalter and Schaffer forthcoming).
- 2 Indeed there is a real debate as to whether evidence must be understood in a truth-conducive way (see, for instance, Fantl and McGrath 2010).
- 3 It is important to note that this discussion references the received view in the epistemic literate on these intuitions, and not experimental evidence directly measuring philosophers' actual judgments. So it's possible that factors like publication bias against those without stakes intuitions could be playing a role in artificially inflating the near consensus about bank cases.
- 4 Not all philosophers report having stakes-sensitive intuitions (see Schaffer 2006). This may point to the existence of important individual differences in bank cases and beyond.
- 5 Another thing that social psychology seems to suggest is that any impact of stakes on knowledge goes through an effect on credence. But if this is the case, then this would show that stakes are not necessarily an independent fourth factor in knowledge along with justified true belief, but merely causally connected to belief (see Weatherson 2005).
- 6 Phelan (forthcoming) has also shown that subject stakes have a marginal impact on people's judgments about evidence. Beyond just looking at stakes and error, researchers have also shown that moral judgment can play a large role in people's willingness to ascribe knowledge (Beebe and Buckwalter 2010; Beebe and Jensen 2012; Buckwalter forthcoming). Presumably, a theory of knowledge

that wished to do justice to folk intuitions would also need to account for these epistemic judgments.

- 7 Marginal results for stakes on ascribers' intuitions in bank cases may also call into question the assumption that the impact stakes have on subject's degrees of confidence mean that stakes will have an impact on ascribers' intuitions.
- 8 Though we might wonder, if the purported effect was driven primarily by accommodation, what would knowledge have to do with it?
- 9 This study used the internet-based commercial research tools Mturk and Qualtrics. Online samples were restricted to participants located in the United States.
- 10 Thirty participants were removed from this study for failure to pass two very basic comprehension check questions.
- 11 Means and standard deviations for denial conditions: Low Error/Low Stakes ($M = 3.48, SD = 1.62$), Low Error/High Stakes ($M = 4.15, SD = 1.05$), High Error/Low Stakes ($M = 4.27, SD = 1.08$), High Error/High Stakes ($M = 3.92, SD = 1.12$). Means and standard deviations for assertion conditions: Low Error/Low Stakes ($M = 4.70, SD = 0.56$), Low Error/High Stakes ($M = 4.48, SD = 0.59$), High Error/Low Stakes ($M = 4.05, SD = 1.30$), High Error/High Stakes ($M = 4.33, SD = 0.73$). In the results to be reported below, a three-way between-subject analysis of variance was conducted to evaluate the effect of error, stakes, and speech-act type on participants' truth-value judgments in the bank cases.
- 12 Main effect for the factor of speech act, ($F(1, 177) = 8.6, p < 0.01$).
- 13 A significant interaction effect was found between factors of speech act and error, ($F(1, 177) = 4.62, p < 0.05$).
- 14 No significant interaction effect was found between speech act and stakes, ($F(1, 177) = 0.40, p = 0.53$).
- 15 Indeed, the only significant effect of stakes was an incredibly complex interaction. In cases with salient error possibilities, people were less inclined to say that the denial was true when the stakes were high ($M = 3.92, SD = 1.12$), whereas in cases without salient error possibilities, people were more inclined to say that the denial was true when the stakes were high ($M = 4.15, SD = 1.05$). In other words, there was an effect such that high stakes had opposite impacts on denials depending on whether error possibilities were made salient ($F(1, 177) = 6.00, p < 0.05$).

- 16 Relative to Lewis (1996), this may suggest that simply mentioning error possibilities is not enough to make them salient in the relevant, epistemic context-altering way.
- 17 For similar stakes results for other verbs besides “know” and “believe,” see Buckwalter and Schaffer forthcoming.
- 18 Ten participants were removed from this study for failure to pass a very basic comprehension check.
- 19 These materials are borrowed directly from Pinillos (2012).
- 20 Typo Low-Stakes Belief ($M = 2.71$, $SD = 1.27$), Typo Low-Stakes Knowledge ($M = 2.61$, $SD = 0.89$), Typo High-Stakes Belief ($M = 6.59$, $SD = 5.05$), Typo High-Stakes Knowledge ($M = 5.12$, $SD = 3.42$). A two-way between-subject analysis of variance was conducted to evaluate the effect of Mental State and Stakes on participant-free responses regarding evidence. A significant main effect was obtained for stakes, $F(1, 86) = 23.1$, $p < 0.01$. However, no main effect was found for Predicate, $F(1, 86) = 1.40$, $p = 0.24$, and no interaction between these two factors was detected, $F(1, 86) = 1.05$, $p = 0.31$.
- 21 The supporter of IRI might respond to this objection by claiming that if knowledge is a norm of belief, then identical results between mental states would be compatible with the impact of stakes on people’s criteria for knowledge ascription. This is certainly a *possibility*, just one that remains to be proven experimentally.
- 22 Indeed, another worry is that a possible ambiguity exists whereby a more natural reading of the question asked in these experiments is something like, “how many times *should* the subject proofread his paper in this situation?”.
- 23 The effect shown here for error is smaller than what was shown in work by Schaffer and Knobe (2012).
- 24 Feltz and Zarpentine investigate a range of life or death cases outside of bank contexts where subject stakes are quite vivid, but are also unable to detect stakes effects.
- 25 Phelan (forthcoming) tests a similar hypothesis by asking participants to judge which factors *should* affect an epistemic subject’s confidence in her beliefs. While participants profess to the general principle that stakes should affect confidence judgments (about evidence at least), they fail to allow the costs of being wrong to influence the *actual* judgments made in these cases when they experience them.
- 26 If this solution is correct, then these differences seem to highlight the need to institute more careful controls when utilizing the evidence-by-intuition method in philosophy. Such methods may be just as susceptible to criticisms one might

- make of any research program in psychology regarding experimental design confounds, or biases (see, e.g., order effects in trolley problem intuitions among professional philosophers by Schwitzgebel and Cushman 2012).
- 27 Specifically, the present evidence may tell against “pragmatist” contextualists, according to which practical matters partly determine the truth of knowledge ascriptions.
- 28 Though it is important to note that this is nonetheless a considerable theoretical blow, since IRI was not on the board before it was claimed that the alleged intuitions needed to be accounted for.
- 29 See, for instance, Weinberg’s notion of “philosophical effect size” whereby the simple detection of a psychological effect may not always be sufficient for supporting certain roles in philosophical argument without meeting a series of further conditions (2011).

References

- Beebe, J. R. and Buckwalter, W. (2010), “The epistemic side-effect effect.” *Mind & Language*, 25, 474–98.
- Beebe, J. R. and Jensen, M. (2012), “Surprising connections between knowledge and action: The robustness of the epistemic side-effect effect.” *Philosophical Psychology*, 25, 689–715.
- Brown, J. (2013), “Experimental philosophy, contextualism and SSI.” *Philosophy and Phenomenological Research*, 86(2), 233–61.
- Buckwalter, W. (2010), “Knowledge isn’t closed on Saturdays.” *Review of Philosophy and Psychology*, 1, 395–406.
- . (2012), “Non-traditional factors in judgments about knowledge.” *Philosophy Compass*, 7, 278–89.
- . (forthcoming), “Gettier made ESEE.” *Philosophical Psychology*.
- Buckwalter, W. and Schaffer, J. (forthcoming), “Knowledge, stakes, and mistakes.” *Noûs*.
- Chatterjee, S., Heath, T. B. and Min, J. (2009), “The susceptibility of mental accounting principles to evaluation mode effects.” *Journal of Behavioral Decision Making*, 22, 120–37.
- Cohen, S. (1988). “How to be a fallibilist.” *Philosophical Perspectives*, 2, 91–123.
- . (1999), “Contextualism, skepticism, and the structure of reasons.” *Philosophical Perspectives*, 13, 57–89.

- . (2004), "Knowledge, assertion, and practical reasoning." *Philosophical Issues*, 14, 482–91.
- DeRose, K. (1992), "Contextualism and knowledge attributions." *Philosophy and Phenomenological Research*, 52, 913–29.
- . (1999), "Contextualism: An explanation and defense," in J. Greco and E. Sosa (eds), *The Blackwell Guide to Epistemology*. Oxford: Basil Blackwell, pp. 182–205.
- . (2005), "The ordinary language basis for contextualism and the new invariantism." *Philosophical Quarterly*, 55, 172–98.
- . (2009), *The Case for Contextualism*. Oxford: Oxford University Press.
- . (2011), "Contextualism, contrastivism, and x-phi surveys." *Philosophical Studies*, 156, 81–110.
- Fantl, J. and McGrath, M. (2002), "Evidence, pragmatics, and justification." *The Philosophical Review*, 111, 67–94.
- . (2010), *Knowledge in an Uncertain World*. Oxford: Oxford University Press.
- Feltz, A. and Zarpentine, C. (2010), "Do you know more when it matters less?" *Philosophical Psychology*, 23, 683–706.
- Fischer, P., Jonas, E., Frey, D. and Kastenmüller, A. (2008), "Selective exposure and decision framing: The impact of gain and loss framing on confirmatory information search after decisions." *Journal of Experimental Social Psychology*, 44, 312–20.
- Hawthorne, J. (2004), *Knowledge and Lotteries*. Oxford: Oxford University Press.
- Hsee, C. K. and Zhang, J. (2004), "Distinction bias: misprediction and mischoice due to joint evaluation." *Journal of Personality and Social Psychology*, 86, 680–95.
- . (2010), "General evaluability theory." *Perspectives on Psychological Science*, 5, 343–55.
- Lewis, D. (1979), "Scorekeeping in a language game." *Journal of Philosophical Logic*, 8, 339–59.
- . (1996), "Elusive knowledge." *Australasian Journal of Philosophy*, 74, 549–67.
- May, J., Sinnott-Armstrong, W., Hull, J. G. and Zimmerman, A. (2010), "Practical interests, relevant alternatives, and knowledge attributions: An empirical study." *Review of Philosophy and Psychology*, 1, 265–73.
- Mayseless, O. and Kruglanski, A. W. (1987), "What makes you so sure? Effects of epistemic motivations on judgmental confidence." *Organizational Behavior and Human Decision Processes*, 39, 162–83.
- Phelan, M. (forthcoming), "Evidence that stakes don't matter for evidence." *Philosophical Psychology*.
- Pinillos, N. Á. (2011), "Some recent work in experimental epistemology." *Philosophy Compass*, 10, 675–88.

- . (2012), "Knowledge, experiments and practical interests," in J. Brown and M. Gerken (eds), *Knowledge Ascriptions*. Oxford: Oxford University Press.
- Schaffer, J. (2006), "The irrelevance of the subject: Against subject-sensitive invariantism." *Philosophical Studies*, 127, 87–107.
- Schaffer, J. and Knobe, J. (2012), "Contrastive knowledge surveyed." *Noûs*, 46(4), 675–708.
- Schwitzgebel, E. and Cushman, F. A. (2012), "Expertise in moral reasoning? Order effects on moral judgment in professional philosophers and non-philosophers." *Mind and Language*, 27(2), 135–53.
- Smith, S. M., Fabrigar, L. R. and Norris, M. E. (2008), "Reflecting on six decades of selective exposure research: Progress, challenges, and opportunities." *Social and Personality Psychology Compass*, 2, 464–93.
- Sripada, C. and Stanley, J. (2012), "Empirical tests of interest-relative invariantism." *Episteme*, 9(1), 3–26.
- Stanley, J. (2005), *Knowledge and Practical Interests*. Oxford: Oxford University Press.
- Weatherson, B. (2005), "Can we do without pragmatic encroachment?" *Philosophical Perspectives*, 19, 417–43.
- Weinberg, J. (2011), "Out of the armchair, and beyond the clipboard: Prospects for the second decade of experimental philosophy." Invited speaker, The 103rd Annual Meeting of the Southern Society for Philosophy and Psychology. March 11, 2011.