9

Knowledge, Experiments, and Practical Interests

Ángel Pinillos

9.1 Introduction

Recently, some philosophers have defended the idea that knowledge is an interest-relative notion. According to this thesis, whether an agent who believes P also counts as knowing P may depend on the practical costs of her being wrong about P. When stakes are high, the epistemic standards for attaining knowledge may be higher. For example, a quick glance at your calendar might suffice to know the time of your haircut appointment, but that quick glance would not suffice if the appointment is a matter of life and death. This perspective marks a radical departure from traditional accounts that take knowledge to be a purely intellectual concept. I think there is much to say on behalf of the interest-relative notion. In this chapter, I report on a series of experimental results which give some support to the thesis that ordinary people's attributions of knowledge are in fact sensitive to practical interests or stakes. Although there has been experimental research showing that decision-making strategies are influenced by stakes (Kahn and Baron 1995; McAllister *et al.* 1979; Mayseles and Kruglanski 1987),[1] to date, there have been no studies which give direct support for the thesis that ordinary attributions of knowledge are sensitive to practical interests in the sense just discussed. On the contrary, some recent empirical work has been thought to disfavor it. Now it is one thing to report on folk behavior and another to argue for a substantive epistemological thesis. As we will see, some philosophers hold that information about the former can inform the latter, and this is the perspective that I take here. Accordingly, I will

I would like to thank G. Shyam Nair and Nick Smith for their invaluable help in this project. I also would like to thank Wesley Buckwalter, Jessica Brown, Craig Enders, Mikkel Gerken, Matt Greve, Joshua Knobe, Ron Mallon, Joshua May, Matthew McGrath, Cecilea Mun, Jennifer Nagel, Shaun Nichols, Jonathan Schaffer, Jason Stanley, Jeff Watson and Jonathan Weinberg. Part of this project was funded in part by a generous grant from the Institute for Humanities Research at Arizona State University. I would also like to thank audiences at the 2010 Eastern APA in Boston, MA, and at the 2011 Western Experimental Philosophy conference at the University of Arizona.

[1] Though by no means should we say that in making decisions when stakes are high, agents always choose strategies which lead to greater accuracy. Perceived high stakes can lead to stress which can in turn result in the adoption of simple heuristics (Luce *et al.* 1997), or it can lead decision makers to focus on features with high affective content (Slovic *et al.*, 2001).

argue that the results I present support an invariantist but interest-relativist account of knowledge, a thesis known as Interest Relative Invariantism (IRI). I do not make the case that IRI gives us the very best explanation of the results. I only want to argue that IRI gives us a simple and elegant explanation of the new data, and that the same cannot be said about traditional contextualism, a leading competitor to IRI. This claim should be qualified right away. It is not that contextualism has great difficulty explaining the data. On the contrary, a lot of the evidence I present is very friendly to contextualism and can even be seen as providing new support for the thesis. Rather, the idea is that the data presented here, on the whole, is better explained by IRI than traditional contextualism.[2]

Before we begin, I wish to briefly point out why I think the type of work presented here may be of some relevance not just to epistemology but also to the cognitive sciences. Philosophers have argued that knowledge has great normative import: Whether an agent knows P will be connected with questions of whether it is acceptable (in a normative sense) for her to use P in reasoning or action including assertion (Hawthorne 2004; Hawthorne and Stanley 2008; Williamson 2000). Some of these ideas are reflected in ordinary usage. Philosophers have observed that people naturally use the concept 'knowledge', for example, to criticize behavior (why didn't you prepare better? you knew the dignitaries would be there!). Arguably, then, knowledge or the concept 'knowledge' plays an important role in our cognitive lives. In fact, according to one major study carried out by Oxford Dictionaries, 'knows' is the 59th most commonly used word of the English language, edging out words like 'take', 'person' and 'good' which also represent important notions.[3] And 'know' is one of the most commonly used mental state words among young children.[4] These considerations suggest that an effort to determine, through controlled experiments, how people actually use the concept knowledge, might be of interest not only to epistemologists but also to researchers interested in how the mind works.[5]

9.2 The relevance of experiments to epistemology

In recent years, two competing theories in epistemology have risen to the main stage: contextualism and Interest Relative Invariantism (IRI). Contextualists hold that different uses of 'A knows that P' can vary in content in a distinctively epistemological

[2] Because of space limitations, I am leaving out discussion on how intellectual invariantism might try to explain the data. This does not mean that I think the theory is less important than IRI or contextualism, or that the data does not have implications for the theory.

[3] Oxford English Corpus: www.oxforddictionaries.com/page/oecfactslanguage/the-oec-facts-about-the-language. Retrieved 02/06/2011.

[4] See, for example, Shatz *et al.* 1983.

[5] There is similarity between the present project and Joshua Knobe's (2010) work. Knobe uses experimental techniques to argue that theory of mind competence is encroached by moral considerations. Similarly, the work here can be interpreted as an argument, using experimental techniques, that competence in knowledge attributions is also encroached by pragmatic considerations. Both theses go against orthodox or purist positions in philosophy. However, it is worth noting that in Pinillos *et al.* (2011), we argue against Knobe's theory.

way (to borrow a phrase from Jason Stanley) where the different contents reflect varying epistemic standards for counting as 'knowing'.[6] Thus, a sentence 'A knows P' might be true if used in a context with lower standards but false when used in a context with higher standards. IRI, on the other hand, denies that contextualism is true and also holds that whether an agent who believes P also knows P at a time may depend on the practical costs of her being wrong about P.[7]

In this chapter, I raise some considerations in favor of IRI (and against contextualism). The central methodological assumption I will be adopting is that information about the behavior and mental states of ordinary people, including careful observation of their deployment of the word 'knowledge', can be relevant in assessing contextualism and IRI. For example, knowledge ascriptions elicited from the folk in certain carefully selected situations can be justifiably taken to express or track true propositions. And this can then be used to inform theories in epistemology.

As it pertains to contextualism, this is not an exotic assumption. Here is Keith DeRose discussing this issue:

> The best grounds for accepting contextualism concerning knowledge attributions come from how knowledge-attributing (and knowledge-denying) sentences are used in ordinary, non-philosophical talk: What ordinary speakers will count as "knowledge" in some non-philosophical contexts they will deny is such in others. (2006: 316)

DeRose's remarks concerning the importance of ordinary linguistic behavior to epistemology should not be surprising. After all, the thesis that a certain word in a language is context sensitive is a thesis that is strongly connected to and is arguably ultimately about the behavior and mental states of competent users of that language. A good example of this methodological assumption at work appears in Ludlow (2005) where evidence for contextualism is gathered by conducting a series of Google internet searches aimed at discovering how ordinary people use the word 'know'. Peter Ludlow's corpus approach, similar in spirit to much experimental philosophy, is perfectly sensible on the common assumption that folk behavior is relevant to the assessment of contextualism.

Now, if folk behavior is relevant to the assessment of contextualism, then it is relevant to the assessment of IRI. This just follows from the fact that IRI entails the

[6] Here is a sample of the vast literature defending contextualism: Cohen (1986, 2005); DeRose (1992, 2006); Heller (1995); Lewis (1996); Ludlow (2005); Neta (2002); Rieber (1998); and Schaffer (2004). There are many ways of developing contextualism. It may be, for example, that the word 'know' behaves like an indexical. Or it may be that there is some other element in the logical form of knowledge ascriptions (like a standard for knowledge or a contrast class) which gets its content fixed by the context of use.

[7] See, for example, Hawthorne (2004); Stanley (2005); Fantl and McGrath (2009); and Weatherson (2011). Some defenders of IRI, like Fantl and McGrath (2009), understand their thesis to be compatible with contextualism. On this hybrid view, 'knowledge' may express different relations (as some contextualists urge), but some of these relations are themselves sensitive to stakes. I will not construe IRI in this way, but nothing much hangs on this. I believe the data presented here does not distinguish between this hybrid interpretation and IRI. But the data supports the hybrid view over traditional contextualism.

denial of contextualism.[8] But just focusing on the stake sensitivity aspect of IRI, it is common for defenders of the thesis to appeal to folk intuitions and behavior. Jason Stanley (2005: 6), for example, thinks that 'our intuitive reactions' to certain vignettes provide 'a prima-facie case for the thesis that knowledge is not just a matter of non-practical facts, but is also a matter of *how much is at stake*'. John Hawthorne (2004: 8), in defending IRI, claims that the inclination to claim that lottery propositions are not known (e.g. the inclination to say that one does not know she will lose the lottery tomorrow) is 'not merely a datum about the inclinations of philosophers'. He also claims the inclinations are 'widespread' and this is a fact that is 'uncontroversial'. Similarly, Hawthorne and Stanley (2008: 571) remark that 'ordinary folk appraisals of the behavior of others suggest that the concept knowledge is intimately intertwined with the rationality of action'. This sentence could be interpreted to mean that they hold that folk judgments can provide some evidence for the intimate connection between knowledge and action which is a core tenet of IRI. It looks then as if some defenders of IRI accept that folk behavior is relevant to establishing their claims.[9]

Now let us ask why folk behavior should be relevant to assessing IRI. According to one possible approach, stake sensitivity is a conceptual or semantic matter. Thus, barring performance error or pragmatic effects, sensitivity to practical interests in knowledge attributions is precisely what we would expect to observe in the behavior of ordinary agents who are competent with 'knows'.[10]

Appealing to semantic or conceptual competence is not the only way to justify the relevance of ordinary people's judgments to epistemology. We could also say that for simple cases, we expect ordinary people to make correct judgments about many mental attitudes, including knowledge. For example, we expect people to not be mistaken about some very simple applications of the concepts of belief, doubt, knowledge, wishing and many others. Furthermore, we might worry that philosophers' judgments concerning some of these simple cases may be biased by theoretical commitment or exposure—giving us a reason to also rely on ordinary people's judgments to get at questions in epistemology.[11]

The previous strategy doesn't assume that philosophers' judgments about certain cases are in fact biased, but it does assume that there is a reasonable suspicion that they

[8] Fantl and McGrath (2009) and Brown (forthcoming) are less optimistic about the relevance of simple first order folk ascriptions of knowledge in reaction to cases for assessing IRI. Brown, in particular, argues against the purported relevance.

[9] Of course, this doesn't mean that all or most arguments for IRI appeal to folk judgments. Fantl and McGrath (2009) and Hawthorne (2004), for example, argue for IRI (or related theses) from certain principles connecting norms and action. Though, see Brown (forthcoming) for how even this strategy could be informed by empirical data from folk intuitions.

[10] Alternatively, we can say that some folk attributions of knowledge are part of common sense and that common-sense judgments are somewhat reliable. This approach is often used in metaphysics. Dean Zimmerman (2007), for example, uses this strategy to advance an A-theory of time.

[11] Some may worry that experimental work may not utilize very simple cases. Progress can be made here by designing experiments which are mindful of existing literature on where humans go wrong in attributing mental states to others. Thanks to Mikkel Gerken here.

are so biased. Is there a reasonable suspicion? I think so. Epistemologists themselves suggest this much. Alvin Goldman (2007) writes:

> philosophers are leery about trusting the intuitions of other philosophical analysts who have promoted general accounts of the analysandum, e.g., knowledge or justification. Commitment to their own favored account can distort their intuitions, even with respect to their own (pre-theoretical) concept. (15)

Related reports of bias appear elsewhere. Szabó Gendler and Hawthorne (2005: 331) give some reason to think that the concept of knowledge might have been 'fashioned and molded by certain philosophical traditions' and consider it valuable to uncover the contours of the "unbiased" concept. Relatedly, Jonathan Schaffer (2006: 90) (who thinks intuitive reactions to cases are relevant to assessing IRI) admits "Perhaps my intuitions are unusual, and no doubt they are theoretically biased." (These remarks should raise eyebrows. If we found analogous reports about some data collection practices in the pages of leading scientific journals, we would have some reason to revisit the methods used for collecting the data.) So given that philosophers themselves admit in print that their judgments or intuitions concerning thought experiments may be subject to bias and distortion, an effort to gather intuitions or judgments from people without those biases may very well be a useful project. This may be an especially prudent course to take when it comes to assessing IRI, since the thesis goes against a long and entrenched "intellectualist" tradition in philosophy. There might be some value then in trying to look at knowledge again with fresh eyes.[12] Appealing to folk behavior can help in this regard.[13]

We just looked at some reasons why folk judgments might be relevant to epistemology. We also also saw that some philosophers accept this perspective. And this is what I will assume for the rest of the chapter.[14] Now, if we assume that folk behavior is relevant to epistemology (in particular, IRI and contextualism), then we should also accept that experimental evidence gathered from folk behavior can also be relevant. For it is hardly credible to hold that although philosophers should pay close attention to the behavior and mental states of ordinary people, we should not do so by using well-established scientific methods including the use of controlled experiments. Thus, I proceed with the methodological assumption that experimental data from the folk can be of relevance to some traditional epistemic concerns.

[12] I am not at all saying that we should abandon philosophers' judgments in this domain. Rather, I am advocating pluralism in methods.

[13] I am assuming that in general, simple folk judgments about knowledge are not biased by commitments or exposure to philosophical theories of knowledge. This seems like a fair assumption. Of course, they could be biased in other ways. Notice, however, that in other domains, folk judgments may be biased in just the ways that would make this sort of strategy problematic.

[14] For more details on the connection between contextualism, IRI and ordinary attributions of knowledge, see Pinillos (forthcoming).

9.3 Evidence-seeking experiments

Recall that according to IRI, whether an agent who believes P also counts as knowing P may depend on the costs of being wrong about P (what is at stake for the agent vis-à-vis P). To illustrate this effect, a philosopher will utilize a pair of nearly identical vignettes except for a difference in what is at stake for the protagonist. If all goes right, we should have the intuition that the protagonist of the vignette knows some proposition in the low stakes case but fails to know a similar proposition in the high stakes case. The asymmetry in our intuitive judgments is then supposed to support the idea that knowledge is sensitive to stakes in the sense of IRI.[15]

There is a body of literature about the epistemic import of these sorts of vignettes (and our responses to them). One common assumption is that the intuitive judgments are in fact asymmetric. A debate exists about which theory, if any, those judgments tended to support.[16] There is certainly not a consensus that these judgments ultimately vindicate IRI over its competitors. IRI, in fact, is very much a minority position in philosophy.[17]

Recently, however, a worry was raised about whether the purported asymmetric judgments to the cases in question really are robust. Three groups of experimental philosophers presented some of the vignettes in question (and others) to non-philosophers.[18] All three research groups essentially settled on the same conclusion: ordinary people's judgments concerning the vignettes did not accord with the judgments philosophers were purportedly sharing about them.[19] This was thought to be problematic not only for IRI but also contextualism (which presumably predicted that ordinary people would respond otherwise) and all the other participants in the debate that assumed the intuitions were robust (even if their contents did not express "true" propositions).[20]

[15] See Stanley (2005), Hawthorne (2004) and the references therein.

[16] Stanley (2005), for example, thinks they support IRI. DeRose (2004) thinks they support contextualism. Nagel (2008), on the other hand holds that the judgments are perfectly consistent with invariantism. Rysiew (2007) thinks they reflect merely pragmatic features of the context.

[17] According to the PhilPapers Survey conducted in 2009 (http://philpapers.org/surveys/results.pl) only 31 percent of philosophers are invariantists. Since IRI is an unorthodox type of invariantism and is new on the scene, I suspect that the percentage of IRI supporters or sympathizers is significantly less than that number.

[18] Buckwalter (forthcoming), Feltz and Zarpentine (forthcoming) and May *et al.* (forthcoming)

[19] Though not every philosopher shared the same judgment. Schaffer (2006) was somewhat skeptical.

[20] What we have here then is a discrepancy between the intuitive judgments of philosophers and the judgments of ordinary people when placed under conditions created by experimental philosophers. This is hardly a new situation. In the last few years experimental philosophers have brought to our attention many cases of this sort. For example, there is literature about how the folk and the philosopher conflict not only about knowledge, but also concerning other central human notions including reference, morality, free will and intentional action (among others). Unfortunately, there is significantly less literature about *why* these differences arise. In Pinillos *et al.* (2011) evidence is presented for the idea that some of the discrepancies (concerning intentional action) begin to disappear if we improve the cognitive or epistemic conditions of the experimental subjects. I do not believe this diagnosis extends to the present case. The 'discrepancies' here, in my view, are simply due to experimental conditions not being appropriate to detect the effect of stakes, and do not reveal genuine differences between the folk and the philosopher.

In this chapter I present the only set of controlled experiments reported thus far in the literature which, taken together, constitute some evidence in favor of IRI. What should we make then of the previous studies which have been taken to disfavor the theory? Perhaps the most we can say at this point is that, on the whole, the experimental evidence is inconclusive about IRI. I do not think this is the right conclusion to draw. First, there are some worries about how damaging the previous results really are for IRI (and contextualism).[21] Second, and putting that aside, it is worth noting that IRI is an existential theory.[22] What this means is that IRI does not predict that every time there is a difference in stakes for agents who believe some true proposition (keeping all else as much as possible constant), there should be a corresponding difference in facts about knowledge. Instead, the theory predicts that a difference in facts about knowledge can sometimes arise because of differences in stakes. So the fact that some experiments fail to detect a stakes effect for a certain class of scenarios while other experiments manage to detect an effect for a very different class of scenarios does not yield, all things being equal, an "inconclusive" verdict. Instead, the totality of the experimental evidence should indeed favor IRI. I believe that this is approximately where we stand now.

With the previous remarks as background, I can now discuss the experimental paradigm, "evidence-seeking" experiments, that I will be developing for the rest of the chapter. As we saw above, in typical thought and controlled experiments, subjects are asked to judge whether a protagonist from a vignette knows some proposition *P* (or asked if and to what extent they agree with some such attribution of knowledge). A challenging feature of this type of experiment is keeping fixed across the various conditions the amount of evidence the protagonist has for proposition *P*. Some commentators have remarked that this is a problematic feature of these experiments.[23] This problem is less severe in evidence-seeking experiments. In these probes, we ask subjects their opinion about how much evidence they think a subject will need to gather before he counts as knowing some proposition. Now if the responses tend to vary as we manipulate the stakes at issue, then this is some evidence that attributions of knowledge are sensitive to stakes in the sense of IRI.

Note now that with these probes, there is less of a problem about keeping the evidence available fixed across the conditions. That feature of the experimental design has now been made a dependent variable which can be directly measured in a way that reflects the stake sensitivity of knowledge ascriptions.

Finally, I should remark that although the experiments below are designed for use in controlled experiments, they can also be used by philosophers themselves to help

[21] Brown (forthcoming), DeRose (forthcoming), Pinillos (forthcoming). See also Schaffer and Knobe (forthcoming) for discussion and further experiments supporting contrastivism (a type of contextualism).

[22] See Weatherson (2011).

[23] See Pinillos (forthcoming) and Jennifer Nagel (forthcoming). For related worries, see DeRose (forthcoming) and Schaffer (2006).

elicit intuitions. As many theorists have remarked to me, the responses detected in the experimental subjects are not very different from the intuitive responses philosophers themselves might give to the scenarios. Thus, the experiments themselves may be useful in a traditional way.

In what follows, I report on experiments I carried out between 2008 and 2011 involving undergraduate students (with little to no exposure to philosophy) studying at the Tempe campus of Arizona State University. The main goal of these studies is to give some evidence that folk attributions of knowledge are sensitive to stakes in the sense of IRI. I call this thesis 'FS–IRI' (Folk Sensitivity–IRI).

9.3.1 Study 1

Method An experiment was conducted to test whether folk attributions of knowledge are sensitive to stakes or practical interest in the sense of IRI (FS–IRI). The sample consisted of undergraduate students at Arizona State University taking introduction level philosophy courses. Assigned by random selection, 77 of them (Low stakes condition) completed the 'Typo-Low' vignette below and 67 of them (High stakes condition) completed the 'Typo-High' vignettes (these vignettes were inspired by G. E. Moore's famous thought experiment concerning typos and common-sense attributions of knowledge).[24,25]

> (Typo-Low): 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): 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.

Now immediately after a subject was given a vignette they were given the following prompt: 'How many times do you think Peter [John] has to proofread his paper before he knows that there are no typos? ____ times.' Subjects were told to insert an appropriate number in the blank space.

[24] See Moore (1939).

[25] These numbers exclude students who failed to follow the instructions.

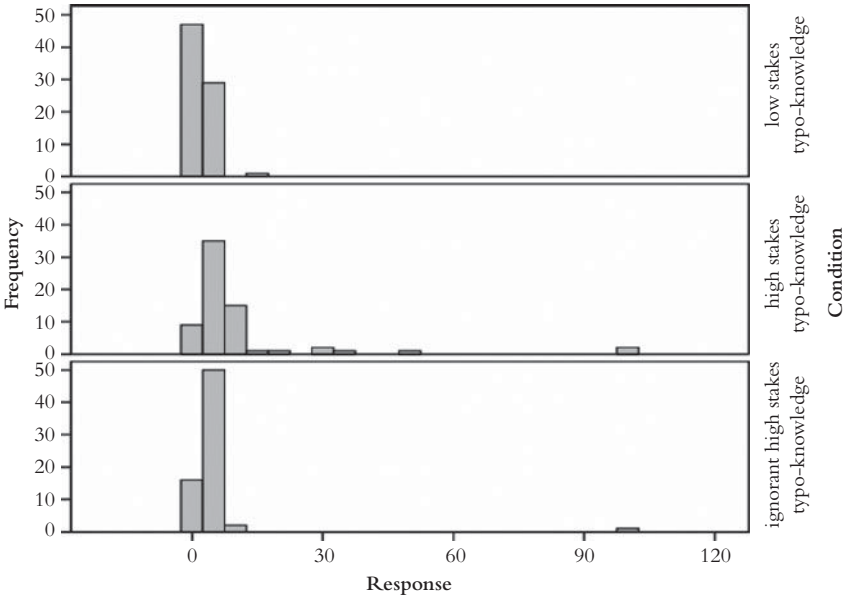

Figure 9.1. Histogram of typo-knowledge (raw data)

Results and discussion In accordance with FS–IRI, we should expect that the number of times a typical subject will think that Peter (Typo-Low) needs to check for typos before he knows there are no typos will differ from the number of times a typical subject will think John (Typo-High) needs to check for typos before he knows there are no typos. It turns out that subjects tended to give higher answers in the high stake condition Typo-High (median=5) than in the low stakes condition Typo-Low (median=2) (see Figure 9.1).[26] The differences in answers are statistically significant to a very high degree and the statistics reveal a large effect (r=–.56).[27] These results then give us some reason to accept FS–IRI: folk attributions of knowledge are indeed sensitive to stakes in the sense of IRI.[28]

[26] Reporting means is not very informative due to the high value of the outliers.

[27] A Mann–Whitney test, appropriate when normality of the data (and equal variance) can be questioned, was used to test for differences between the groups. The test reveals that there is a statistically significant difference between the Typo-Low (median=2) and Typo-High (median=5) groups. Mann–Whitney (N=144) U = 920.500, z=-6.786, r=–.56, p<.001.

[28] A worry here is that the word count between the probes are dissimilar (creating cognitive load associated confounds). Efforts to make the word counts congruent may, however, give rise to more serious problems (either not making the high stakes believable or adding extraneous information to the typo-low probe). Jonathan Schaffer, however, reported to me that he ran similar probes where word counts across the vignettes were similar and he found essentially the same result that I am reporting. I have since ran several evidence-seeking experiments controlling for word count and found similar results.

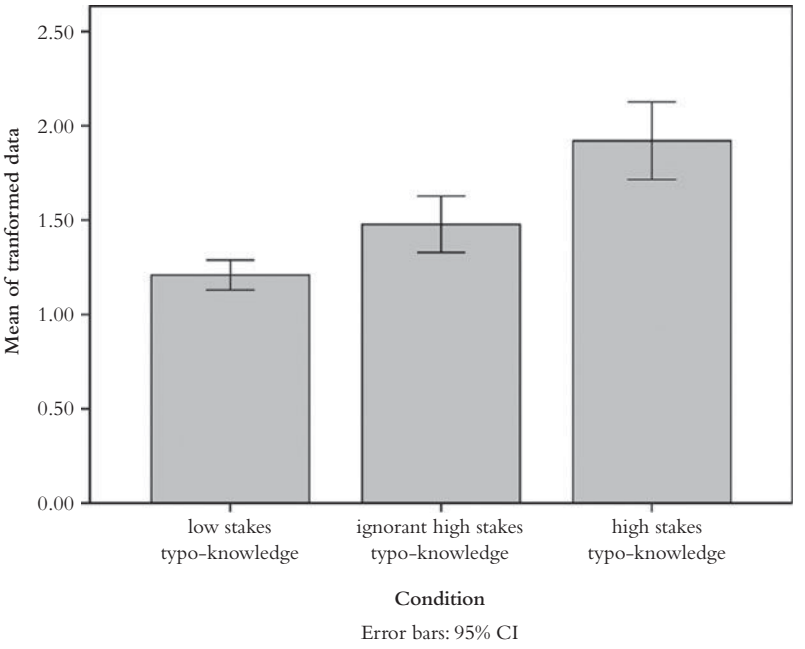

Figure 9.2. Means of typo-knowledge (transformed)

Several points are worth highlighting (the articulation of some of these points report on follow-up experiments). (i) Since data collected in these experiments (see Figure 9.1) do not satisfy very well the assumptions of parametric testing, I have used non-parametric methods throughout. However, in addition, I will often supply parametric statistics performed on a transformation of the data (see Figure 9.2). The transformed data better satisfies the assumptions of parametric statistics.[29] (ii) Since 'evidence-seeking' thought experiments are original with this study, it is important that this type of finding be replicated. I report that it has been replicated by other researchers[30] and in other experiments which I discuss below.[31] (iii) In other variations of this probe involving a different theme (see the 'counting contest case' in the Appendix), I measured

[29] I used a common formula: ln(x+1).

[30] Wesley Buckwalter: http://experimentalphilosophy.typepad.com/experimental_philosophy/2010/07/stakes-and-evidenceseeking.html; also, Jonathan Schaffer (personal communication).

[31] In the spirit of full disclosure, two pilots were ran where the predicted asymmetries were not found to be statistically significant. I do not discuss these here because I believe they were seriously flawed. First, they both concerned medical experts performing when lives are at stake (high stakes). I think the problem was that subjects tend to think that the bar for knowledge is set low here precisely because the experts are trained to perform when lives are at stake. Second, one of the pilots contained an egregious error in the setting up of the stakes.

the cognitive reflectiveness of participants using the Cognitive Reflection Test (CRT).[32] Scoring higher on the test was not correlated with higher or lower responses to the vignettes.[33] This should somewhat alleviate worries that the answers reported here were due to participants not carefully considering the questions. (iv) It might be thought that the results can be explained by appealing to the notion that belief is sensitive to stakes. The idea here is that the respondents think it takes more evidence to know in high stakes because belief is a necessary condition for knowledge and an agent aware of the high stakes will seek more evidence before he forms the relevant belief (and hence come to have knowledge). I have three responses to this suggestion.[34]

(1) Note that appealing to the stake sensitivity of belief does not give us a full explanation of the results. After all, the prompt was about knowledge and not belief. Let us grant that many participants thought that low stakes Peter will need to proofread twice before he believes there are no typos while high stakes John will need to proofread five times before he believes there are no typos. But why don't participants think that low stakes Peter needs to check additional times before he goes from merely believing to knowing? It is, in fact consistent with the stake sensitivity of belief that participants would give the same answers for low and high stakes in our "knowledge" prompts. But they didn't. This needs explaining.[35]

(2) To help alleviate this worry, I ran an "ignorant" high stakes case (Ignorant-Typo-High) (N=69). This is a probe that is just like the Typo-High except that the agent is not aware of the high stakes at issue (see Appendix). Answers to that vignette were different from the answers to the regular low stakes case (Ignorant-Typo-High median=3, Typo-Low median=2). The difference is statistically significant and the statistics reveal a medium size effect (r=–.39).[36] But now

[32] See Frederick (2005). The CRT is highly correlated with putative measures of general intelligence including the SAT verbal, SAT math, Need for Cognition test and the WPT.

[33] The Kruskal–Wallis test used to compare CRT scores (0, 1, 2, 3) in the low stakes condition (N=166) reveals there is no statistically significant difference in responses between the groups χ^2 (3) = 4.364, p=.225. The same conclusion holds for the high stakes condition (N=79): χ^2 (3) = .591 and p=.898.

[34] I would like to thank Wesley Buckwalter, Joshua May and Jennifer Nagel for pressing me on this point.

[35] Mikkel Gerken suggests that subjects may hold that all conditions for knowledge have been met in both cases except for belief in the high stakes case. Depending on how this suggestion is cashed out, it may have the consequence that if the high stake protagonist were simply to form the relevant belief after proofreading twice (or some low number), then he would have knowledge. What needs to be explained now is this: why doesn't the protagonist come to have the belief, as it were, after proofreading just twice? After all, he would know by simply forming the belief.

[36] A Kruskal–Wallis test H(2)=53.23, p<.001 (non-parametric version of a one way ANOVA) shows that there are differences between the three groups (Typo-Low, Typo-High and Ignorant-Typo-High). A Mann–Whitney test, appropriate when normality of the data (and equal variance) can be questioned, was used for pair-wise comparisons. The test reveals that there is a statistically significant difference between the Typo-Low (median=2) and Typo-High (median=5) groups (as reported above). Mann–Whitney (N=144) U = 920.500, z=-6.786, r=–.56, p<.001. A similar thing can be said about the comparison between Typo-Low and Ignorant-Typo-High: Mann–Whitney (N=146) U=1497.000, z=-4.693, r=–.39, p<.001 (Bonferroni corrections applied to set the threshold for significance). In addition, a one way between subjects

these facts cannot be straightforwardly explained by the idea we are considering. This is because since the protagonist is not aware of the high stakes, he is not going to check more times before he forms the relevant belief.

(3) I ran another probe to rule out the explanation we are considering. Here, I stipulated that the protagonist formed the belief that there were no typos soon after he finished writing his paper. So it is now less likely that the responses reflect the idea that the high stakes protagonist collects more evidence before forming a belief. In this new probe, subjects were presented with either the Typo-High or Typo-Low vignettes, but this time the vignettes were followed by this prompt (after being given two reading comprehension prompts):

It turns out that right after Peter/John finished writing his paper, he formed the belief that there are no typos in his paper—and in fact there are no typos. But does he know this? How many times do you think that Peter/John has to proofread his paper before he knows there are no typos? ___ times.

The results of this study reveal that folk use of knowledge is sensitive to stakes even when we attempt to fix belief. There was a (highly) statistically significant difference in responses between the high (Belief-Typo-High) and low stake cases (Belief-Typo-Low). This statistic also revealed a medium size effect (r=.34).[37] This data then puts some pressure against the objection we are considering.

This concludes the responses to the objection that the results are explained by appealing to the stake sensitivity of belief.

Finally, I will consider whether the subjects are misunderstanding what I was trying to ask them. It might be that subjects are understanding the prompt about knowledge as a prompt asking a normative question which is not about knowledge ("how many times should Peter/John proofread his paper before he turns in his final draft?") *and* that this alternative understanding is either due to a performance error or perhaps due to a triggering of a conversational implicature.[38] In response to this sort of worry, I ran a pair of prompts (low stakes and ignorant high stakes). The difference between these and the previous ones is (i) the existence of a leading prompt which is just the

ANOVA was performed on the natural log transformation of the data discussed earlier. Using this method reveals that stakes play a role in knowledge attributions. $F(2,201)=23.33$, $p.<001$, $\eta^2 = .18$. A post hoc analysis using the Games-Howell procedure also reveals significant differences between Typo-Low (M=1.2, SD=.35, 95% *CI* [1.13, 1.28]) and Typo-High (M=1.9, SD=.84, 95% *CI* [1.72. 2.13]), p<.001. There is also a difference between Typo-Low and Ignorant-Typo-High (M=1.47, SD=.62, 95% [1.33,1.66]), p=.006. And, in addition, between Typo-High and Ignorant-Typo-High, p=.002. I discuss this last comparison below.

[37] Median for Belief-Typo-Low here was 2 and for Belief-Typo-High is 3. Mann–Whitney (N=78) Z=-2.98, r=.34, U=472.500, p=.003

[38] As I go on to discuss in study 3, I do think it is plausible that many subjects are replacing the 'knowledge' question with the normative one. But I think this is no error on their part. On some versions of IRI, those questions are actually equivalent!

normative question from above ('John/Peter should proofread his paper___times before he turns it in'), and (ii) a parenthetical remark following the standard knowledge prompt to the effect that the subject should recognize that they are being asked about 'knowledge' not about some other notion ('NOTE THE DIFFERENCE BETWEEN THIS QUESTION AND THE PREVIOUS ONE. ONLY THE SECOND ONE CONTAINS THE WORD "KNOWS"'). The features of these modified stimuli should then make it less likely that subjects are making a performance error or attending to a pragmatic implicature. It turns out that neither the low stakes nor the high stakes responses were affected by having this extra clarifying material added to the stimuli.[39]

Before I move on I will discuss briefly the fact that answers to Ignorant-typo-High (median=3) and Typo-High (median=5) were different. Since the stakes are the same for both (high), the difference in answers were likely not due to perceived stakes (on the part of subjects). The relevant difference between these probes is that in one but not the other condition, the protagonist is aware of the stakes involved. This difference will probably affect the level of confidence subjects will perceive our protagonist to possess. It is plausible to think that subjects accept that confidence plays a role in knowledge and this idea could help explain this statistic.[40]

9.3.2 Study 2

Here, we tested to see whether FS–IRI holds under further reflection. Subjects read both high and low scenarios before they are asked to give their answers to each. I believe that this is a particularly important probe to run for this case (as well as for many philosophically relevant thought experiments). The way pairs of thought experiments are presented in the philosophy literature, the reader finds herself being aware of both vignettes before settling on intuitions. Thus, by running this type of 'juxtaposed' experiment, we place subjects in conditions that are closer to the conditions philosophers experience when they react to these cases. Moreover, there is reason to think that by allowing subjects to compare and contrast both scenarios before they give answers to each, they would be in a better position to detect a perceived error in their initial

[39] For low stakes: N=75, Mann–Whitney U=545.500, Z=-.926, p=.34, r=-.11. For high stakes: N=99, Mann–Whitney U=1058.500, Z=-1.16, p=.26, r=-.12.

[40] The raw data includes three values that exceeded 50 and should be considered extreme. Three subjects wrote down exactly '100' for their answer (two in Typo-High and one in Ignorant-Typo-High). It is likely that these subjects were thinking about the stimulus in a very special way. It might be that either the mentioning of high stakes put them in a certain skeptical or philosophical mindset, or perhaps these subjects are simply skeptics about knowledge and the round "100" answer is an expression of their skepticism (or perhaps the response is flippant). An argument can be made that those numbers should be left out of the data analysis. I did not exclude them. But a one-way ANOVA analysis on the raw numbers, deleting those three responses, yield essentially the same results as above. F(2, 207)=17.16, p<.001. A Games-Howell post hoc procedure reveals significant differences between Typo-Low (M=2.61, SD=1.96, 95%CI[2.17, 3.06]) and Typo-High (M=7.38, SD=8.55, 95%CI[5.27, 9.5]), p<.001.There is also a significant difference between Typo-Low and Ignorant-Typo-High (M=3.63, SD=1.95, 95% CI[3.16, 4.1]), p=.006.

inclination to give different answers.[41] Hence, if we find that subjects stick to their 'stake sensitive' answers, this is some evidence that they do not perceive stake sensitivity (in knowledge ascriptions) to be the product of error. This would not only be evidence for FS–IRI, but it is also evidence that people's attributions of knowledge are sensitive to stakes even under further reflection.[42]

Method The sample consisted of 95 students taking introductory courses at Arizona State University. The students were given a survey sheet that included both the Typo-High and the Typo-Low vignettes. The subjects were instructed that the vignettes were unrelated and that they were not to write their answers at the end of the sheet until they read both vignettes. The following was found at the bottom of the sheet: 'Peter needs to proofread his paper ___ times before he knows there are no typos, and John needs to proofread his paper ___ times before he knows there are no typos.' The subjects were asked to fill in both blanks with numbers. The ordering of the vignettes was counter-balanced as was the ordering within the juxtaposed prompt (Peter or John first).[43]

Results and discussion As predicted by FS–IRI, subjects' answers for the low stakes (median=2) part of the prompt (Peter) differed from those for high stakes (median=3) part of the prompt (John).[44] The difference in these answers was (highly) statistically significant.[45] In addition, A paired t-test was performed to measure the mean of the differences on the transformed data. The mean of the differences (M=.54, SD=.69, N=95) was significantly greater than zero, $t(94)$=-7.6, p<.001. The effect size is medium but close to large (d=.78). These results support FS–IRI.

More can be learned from this experiment.[46] The percentage of subjects who gave asymmetric answers in the direction consistent with IRI (a higher number for high

[41] The conjunction fallacy, for example, can be eliminated by a similar method. See Kahneman and Tversky (1996). Pinillos *et al.* (2011) use a similar method to show that the Knobe effect can be reduced in this type of condition.

[42] Of course, the conditions of this experiment are are not identical to ones that philosophers enjoy. Philosophers might note the problem with the juxtaposed cases only when they recognize that it conflicts with a principle about the irrelevance of practical interests to knowledge. This is something we might not expect ordinary people to do. Thanks to Matthew McGrath here.

[43] Ordering had no detectable effect on the answers for the high stakes responses, but there was some effect (which can be questioned, see below) for the low stakes answers. A Kruskal–Wallis test was used to compare the four orderings for ordering effects. There were no statistically significant differences found for the high stakes answers. N=95, $\chi^2(3)$ = .434 and p=.933. There was, however, an ordering effect for the low stakes responses N=95, $\chi^2(3)$=14.809 and p=.002. But the effects were only detected in two of the six comparisons (at significance set at .05/6). They occurred between the low-stakes-first-vignette/low-stakes-first-juxtaposed-prompt and low-stakes-first-vignette/high-stakes-first-juxtaposed-prompt conditions, as well as between the former and the high-stakes-first-in-the-vignette/low-stakes-first-in-the-prompt conditions. These results, however, do not reveal a distinct pattern or a theoretically interesting effect. A further replication of this experiment (not included in this data) with a larger sample reveals no ordering effects for either the high stakes or low stakes answers. This leads me to believe that if there are ordering effects for the low stakes answer, it is not a very pronounced effect—but may be worth exploring further.

[44] Answers that did not specify a unique number were discarded.

[45] A Wilcoxon signed-rank test was performed, N=95, p<.001.

[46] For this analysis, only surveys with both blanks filled with amounts (including ranges) were counted.

stakes) is 68 percent. This number is higher than would be predicted by chance (or at most 50 percent). This is a statistically significant difference.[47] These results further bolster the case for FS–IRI.[48]

As I mentioned earlier, the juxtaposed cases allow subjects to compare their own answers. They then have a chance to correct them if they sense that they are mistaken. The evidence suggests, however, that subjects do not perceive that giving asymmetric answers is a product of error. It seems then that folk attributions of knowledge are sensitive to stakes even under further reflection. This constitutes more evidence for FS–IRI.

9.3.3 Study 3

The idea that folk attributions of knowledge are sensitive to stakes in the sense of IRI (FS–IRI) is closely related to the claim that ordinary people use knowledge claims to appraise behavior. In this experiment I test for these folk appraisals and take the results to be highly relevant to assessing FS–IRI. Before I get to that, I need to cover some preliminaries.

Preliminaries Many philosophers have argued that there is an intimate connection between knowledge and various norms of behavior. One formulation of this theme recently defended by John Hawthorne and Jason Stanley (2008) is dubbed the Reason–Knowledge Principle (RKP).

> RKP: Where one's choice is p-dependent, it is appropriate to treat the proposition that p as a reason for acting if and only if you know that p.[49]

One key piece of evidence for this principle is the idea that the concept 'knowledge' is often used in appraising action. In particular, it is very natural to criticize someone's use of a premise in reasoning by claiming they do not know that premise. Suppose, for example, that I decline health insurance coverage for the summer. I may reason that since I won't get sick, I won't get health insurance. A natural way to criticize me is by claiming that I should get insurance since I do not know I won't get sick. Similarly, if I sell a lottery ticket I bought for a dollar for 1 cent citing the premise that I will lose, you

[47] χ^2 goodness-of-fit (N=97) = 11.23, $p < .001$

[48] It might be thought that the results of the juxtaposed experiment reveal a reduction of the effect of stakes. I do not believe that we have enough information to say this. The fact that the median response to the high stakes prompt (in the juxtaposed case) is lower than the median response in the original high stakes study only reveals that the assessment of the evidence required for knowledge may be affected by exposure to similar cases. A reduction of the effect, in my view, would be found by comparing the percentage of people who make asymmetric judgments when not being exposed to the juxtaposed case with the percentage of people who accept the asymmetric judgments while being exposed to the juxtaposed cases. The former percentage is not available nor can it be recovered from the existing data.

[49] Fantl and McGrath (2007) defend only a weaker principle: 'S knows that p only if S is rational to act as if p'. Many philosophers have argued against principles connecting knowledge and action. See, for example, Brown (2008) and Neta (2007).

may criticize me by saying that I don't know that I will lose the lottery (Hawthorne and Stanley 2008; Hawthorne 2004).

Now, here are two relevant points about RKP. First, as the discussion in the previous paragraph illustrates, support for RKP comes from (among other things) ordinary practices of behavior appraisal. Hawthorne and Stanley (2008: 571), for example, motivate the principle in part by appealing to ordinary linguistic behavior. They are upfront about this when they say: 'ordinary folk appraisals of the behavior of others suggests that the concept of knowledge is intimately intertwined with the rationality of action'. I do think that it is plausible that ordinary people implicitly accept RKP. The second point I would like to make about RKP is that it if it is true, then knowledge is sensitive to stakes (assuming fallibilism).[50]

Let us put these two points together. If ordinary people implicitly accept RKP, and if RKP entails that knowledge is sensitive to stakes, then this strongly suggests that ordinary people behave as if knowledge is sensitive to stakes in the intended sense. That is, FS–IRI is true. Of course, there is no strict entailment here. But it would be very surprising if it turned out that ordinary people appraise behavior with knowledge ascriptions in accordance to RKP, yet their knowledge ascriptions were not sensitive to stakes.[51,52]

Indirect probes to test FS–IRI Here is then how we can turn this idea to gather further evidence for FS–IRI: If it turns out that the folk implicitly accept RKP, then this will be further evidence that FS–IRI is true. Suppose we were to find that when ordinary people were asked if X knows P (in some carefully selected situation), then they would often answer a different question: whether it is appropriate for X to use P as a reason for acting. If we found that people did this, then this would be some evidence that they were behaving as if those questions were equivalent, and so this would be some evidence that they implicitly accept RKP.[53]

[50] Let A and B both have the same evidentiary support for and both believe P. A knows P but has very little at stake in P. In accordance to RKP, it is appropriate for A to use P in her reasoning. But suppose B has a lot at stake concerning P so that it would not be appropriate for B to act on P unless she gathered even more evidence for it (fallibilism assumed). Now according to RKP, B would not know P in this instance. So knowledge is sensitive to stakes. See Fantl and McGrath (2007, 2009) for more details on this type of argumentation. But see DeRose (2009) for an account of the connection between action and knowledge that is compatible with contextualism.

[51] I am not suggesting here that people will accept the proposition that knowledge is sensitive to stakes on account of accepting RKP or even that they accept that (former) proposition. In general, we don't expect people to believe the consequences of all the principles they accept. What I am suggesting is that implicit acceptance of a principle connecting knowledge and action would show up in some patterns of knowledge attributions. The general idea is natural. For example, if people accepted some principle connecting knowledge and truth (knowledge requires truth) or justification (knowledge requires justification), we would be surprised to learn that acceptance of these principles did not show up in some of their first order attributions of knowledge. Thanks to Jessica Brown here.

[52] See also Gerken (ms).

[53] The idea that subjects replace one question for another is endorsed by Kahneman and Frederick (2002: 53) in some cases where one question is a heuristic short-cut for another: 'when confronted with a difficult question people often answer an easier one instead, usually without being aware of the substitution'. I discuss below the possibility that RKP is associated with a heuristic.

Preliminary evidence Let us turn to the surveys. In the juxtaposed case from experiment 2, subjects were given the option of explaining their answers. That is, they were asked to explain why they think that 'Peter needs to proofread his paper X times before he knows there are no typos, and John needs to proofread his paper Y times before he knows there are no typos' (where they filled in X and Y).[54] Out of the 66 subjects who gave asymmetric responses consistent with stakes having the predicted effect, 46 chose to explain their answers. What was found is truly surprising. Although the prompts were clearly about knowledge, only one person in this group actually used the word-form 'know' in their answer. Let us look at this more closely. Out of that group of 46, one type of answer was salient. 25 of them explained their answers by claiming some variation on the theme that John is required to/needs to/should check more times because the stakes are higher for him (no other more dominating pattern of answers were found except those that simply mention "stakes"). These answers reflect the simple idea that the normative requirements for Peter and John differ because of the differences in stakes.[55]

I believe that a good way to explain these observations is by appealing to the notion that many subjects in the experiment implicitly accept RKP. To see this in more detail, consider what RKP would predict about the prompt 'Peter/John needs to proofread his paper X/Y times before he knows there are no typos'. According to RKP, this statement is equivalent to 'Peter/John needs to proofread his paper X/Y times before it is appropriate for him to treat the proposition that there are no typos as a reason for acting'. This statement is somewhat wordy. But given the context of the vignettes (in which it is naturally presupposed that the protagonists will turn their papers in to the teacher), the following is a natural paraphrase: 'Peter/John should proofread his paper at least X/Y times before he turns it in'. Now suppose that when subjects were asked to give their answers to the original prompt they considered instead this paraphrase. This means they would interpret the original prompt as (⋆)

(⋆) Peter (low) should proofread his paper at least X times before he turns it in and John (high) should proofread his paper at least Y times before he turns it in.

Now, a perfectly natural explanation of why one thinks (in effect) that X<Y, is that since the stakes are different, John should proofread more times than Peter. And this is precisely the sort of explanation our group of 25 gave. Furthermore, this is a response that does not at all invoke the concept 'knowledge' which is also consistent with the discovery that only one subject appealed to that word in explaining her answers. It looks then as if our explanation is on the right track.[56]

[54] Again, the order of 'John' and 'Peter' in the prompt was counterbalanced.

[55] The responses were interpreted by a coder not aware of the goal of the study.

[56] Matthew McGrath points out to me that the paraphrase given (⋆) is more plausible for the high stakes case than the low stakes case. If it only works for the high stakes case, the evidence would still support the hypothesis though in a weakened form.

Let us recap. It was discovered that the explanations that subjects gave to the juxtaposed prompts were highly puzzling because subjects did not at all invoke the word-form 'know'. Moreover, about half of the subjects that displayed stake sensitivity to knowledge ascriptions (as revealed by their numerical answers) and gave explanations said that John should do more than Peter because the stakes are higher for him (or something similar). One plausible account of these puzzling results is that agents are interpreting the original prompt as (⋆). And a good explanation for why they are doing this is that they implicitly accept RKP. This is because according to RKP, the original prompt is equivalent (or very nearly so) to (⋆). Finally, as I mentioned before, if people behave as if RKP is true, then this is good evidence that ordinary people's knowledge ascriptions are sensitive to stakes. This is just FS–IRI.

So far, so good. It looks like we have collected a small amount of evidence that people implicitly accept RKP and hence FS–IRI is true. Here's one more experiment that further supports this idea.

Method An experiment was conducted to see if people implicitly accept RKP. For this study, I devised a pair of evidence-seeking vignettes that are similar to Typo-High and Typo-Low.[57] The vignettes concern a protagonist, Peter, who must correctly count the pennies in a medium-sized jar to win a prize. In the high stakes condition, Count-High, if he gets the count right he wins a lot of money that he can use to save his mother's life. In the low stakes conditions, Count-Low, he wins a pair of movie passes. Subjects were asked to fill in the blank in the following 'knowledge' statement: 'Peter must count the pennies in the jar at least ___ times before he comes to know the correct answer'. As with the typo cases and in accordance with FS–IRI (and our previous results), we should expect subjects to write in a higher number in the high stakes vignette.

(*N*=) 263 undergraduate students taking introduction to philosophy courses at Arizona State University were randomly assigned either the Count-Low or Count-High vignettes followed by the 'Knowledge' prompt above or followed by a 'Normative' prompt to be explained below (yielding four groups in total). Let us call the two groups with the knowledge prompt, 'Count-Low-Knowledge' and 'Count-High-Knowledge'.[58] The normative prompt involved this 'normative' statement: 'Peter should count the pennies in the jar at least___ times before turning in his final answer'. Let us call these two normative groups 'Count-Low-Normative' and 'Count-High-Normative'.[59] Now assuming RKP and given the context of the vignettes, the

[57] I consider different vignettes because I am now reporting on an experiment I ran a few months before I developed the Typo-High and Typo-Low vignettes. See the Appendix for the full vignettes.

[58] Just as with the study 1, there was a statistically significant difference between these numbers, which further supports FS–IRI and is a further replication of our 'Evidence-Seeking' thought experiments. Mann–Whitney (N=128) U=1124.500, p<.001, r=–.4. The median answer for Count-Low-Knowledge was 2 and the median for Count-High-Knowledge was 3.5.

[59] The answers these two groups gave were, of course, different and this difference was statistically significant. Mann–Whitney (N=125) U=798.500, p<.01. The median for Count-Low Normative=2, and for Count-High-Normative=5.

'knowledge' and 'normative' prompts are equivalent (or very nearly so). So if in fact people implicitly accept RKP, for each level of stakes (low or high) the answers to the knowledge prompts should not differ from the answers to the normative prompts. Furthermore, there should not be any interaction effects among the "stakes" and the "knowledge/normative" factor.

Results and discussion In accordance with the hypothesis that people implicitly accept RKP, there was no statistically significant difference between the 'Count-High-Knowledge' answers and the 'Count-High-Normative' answers.[60] Furthermore, there was no statistically significant difference between the answers to the 'Count-Low-Knowledge' answers and the 'Count-Low-Normative' answers.[61] Also, the effects, *r*, associated with these statistics are very small. Using these measures, we are unable to disprove the hypothesis that people implicitly accept RKP. Furthermore, a 2 x 2 between subjects ANOVA was conducted with transformed responses as the

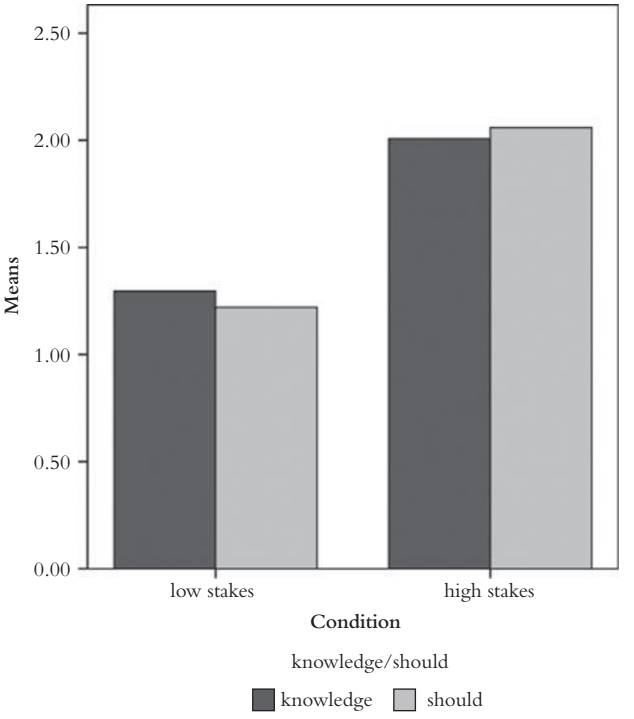

Figure 9.3. Means of counting contest (transformed)

[60] Mann–Whitney (*N*=128) *U*=1853.500, *Z*=-.938, *p*=.348, *r*=-.08.
[61] Mann–Whitney (*N*=125) *U*=1928.500, *Z*=-.12, *p*=.904, *r*=-.01.

dependent variable and stakes (low/high) and "attitude" (knowledge/normative) as independent variables (see Figure 9.3). There was a significant main effect for stakes (as expected), $F(1, 249)=41.39$, $p<.001$, partial $\eta^2=.142$. But there was no significant main effect for "attitude" (knowledge vs. normative), $F(1, 249)=.01$, $p=.92$, partial $\eta^2>.001$. And no interaction effect between the two factors $F(1, 249)=.276$, $p=.6$, partial $\eta^2=.001$. Furthermore, the effect sizes, η^2, for the last two statistics were very small. These results give further indirect support for the thesis that people implicitly accept RKP since the results indicate that people are treating the knowledge question and the "normative" question in similar ways.[62]

These results taken on their own do not come close to *showing* that people behave as if RKP is true. But the idea gains plausibility when it is combined with (a) the data and discussion about the written explanations subjects gave to the juxtaposed vignettes, and (b) the philosophical reasons for RKP including anecdotal evidence about its connection to ordinary appraisals of behavior. And as I mentioned above, that people implicitly accept RKP supports FS–IRI.

At this point, it is important to note that the claim I am making (that people implicitly accept RKP) can be interpreted in a variety of ways. The most straightforward way is one in which we simply say that people (implicitly) believe the principle. And given the remarks at the beginning of this chapter, this would be some evidence for RKP and IRI.[63] However, the claim that people implicitly accept RKP may be thought to be compatible with the idea that the principle provides a useful heuristic in the following sense: Sometimes, when attempting to answer a hard question about knowledge, people replace that question with an easier one about norms of behavior or reasoning (in accordance with RKP). This weaker interpretation is then compatible with the claim that people do not believe that RKP is true. Nonetheless, even if RKP provides a heuristic in this way, it would still be the case that folk attributions of knowledge are sensitive to stakes in some sense. However, it would now not at all be clear that we still have support for IRI. The answer might depend on how heuristics in cognition connect up to conceptual competence.[64] The issues here are complex and uncertain, and for this reason caution should be used in interpreting this study. Nonetheless, I think it is still fair to say that we have gathered some new, if speculative,

[62] Jonathan Schaffer and Wesley Buckwalter have run evidence-seeking experiments replacing 'knows' with 'belief' and 'guess'. They have failed to find differences between responses to these alternate questions and the original with 'knows'. It is not clear to me how this will impact the interpretation of my data. Much will depend on how people understand the various norms for belief and assertions (including assertions that one guesses that P). I should also point out that in a further set of studies, I replaced the evidence-seeking questions with standard queries about the level of agreement with simple knowledge statements 'X knows P'. I found evidence of stake sensitivity here as well. This is evidence that the evidence-seeking probes are tracking knowledge contrary to what the Schaffer–Buckwalter findings might suggest.

[63] Jessica Brown (2008) has produced intuitive counter-examples to RKP and similar principles. If she is right, then perhaps RKP is "aliefed" in the sense of Szabó Gendler (2008) (directing "belief-discordant" behavior) or is a heuristic as I explain below. And now it might be difficult to ascertain whether we really have found support for IRI with this experiment.

[64] For a discussion on this topic, see Gerken (this volume).

indirect evidence for FS–IRI and IRI. For example, it could have turned out that responses to the normative and knowledge prompts were very different, but it didn't turn out that way. Certainly, more work needs to be done in this direction.

9.4 General discussion

I believe that the results of these three studies together support FS–IRI, the claim that folk attributions of knowledge are sensitive to stakes in the sense of IRI. In the first study, there was a statistically significant difference between the amount of evidence people are willing to say is required for knowledge in high and low stakes cases even for reflective individuals (also between the 'ignorant' high stakes case and the low stakes case). In the second study, a similar difference was found even when the cases are juxtaposed. This indicates that attributions of knowledge are sensitive to stakes even under further reflection. The data associated with the third study also support FS–IRI. There, it was found that people's explanations of their answers to the juxtaposed prompts are well accounted for by the idea that they implicitly accept RKP. Furthermore, the fact that there is no significant difference between subjects' responses to certain 'knowledge' and 'normative' queries is also explained by making the same assumption. And, as I discussed above, these results further support FS–IRI. Moreover, when we combine these experimental results with the fact that many philosophers report intuitions in accordance with FS–IRI, we have some good reasons to accept FS–IRI.

9.5 IRI gives an elegant explanation of the results

In this chapter, I have provided some new evidence for FS–IRI. These results also support IRI. The latter move is justified if we make the plausible assumption that people's responses to our simple probes are accurate (see section 9.2 and Pinillos forthcoming for details on the connection between ordinary judgments and IRI). On this construal, there is no need to attribute to subjects performance errors, and there is no need to invoke complicated pragmatic effects. In the first study, differences in answers between Typo-Low and Typo-High and also between Typo-Low and Ignorant-Typo-High are cleanly explained by IRI. This is because according to IRI, it takes more to know when stakes are higher. In the second study, these differences persisted even when the vignettes and prompts were juxtaposed giving subjects a chance to note the differences in their answers and possibly correct them. Hence, even knowledge attributions made in a more reflective state are neatly explained by IRI. Finally, in the third study we saw evidence that people implicitly accept RKP. A simple explanation of this is that RKP is in fact true.[65] But now since RKP entails (assuming fallibilism) that knowledge is sensitive to stakes, the results of this study are also explained by IRI.

[65] I do not mean to say that in every case in which an agent implicitly holds P, the best explanation invokes that P is true. However, in some simple cases, this seems right. For example, if we discover that

All of these considerations, jointly, support the idea that IRI gives us an elegant explanation of the results presented in this chapter.

9.6 Results favor IRI over contextualism

Not only do I think that IRI gives a good explanation of the results presented here but I also believe that contextualism will fare worse with respect to the data. I now turn to this.[66]

As explained earlier, a contextualist about knowledge holds that uses of a single sentence 'A knows P' can vary in content (and truth-value) in epistemically interesting ways. These different contents reflect different epistemic standards for 'knowing'. Defenders of IRI (construed here as incompatible with contextualism), on the other hand, deny this, though they believe that the cost of being wrong about P at a time (stakes) can affect how much evidence a person who already believes P needs to have before she counts as knowing P at the time. In the case of contextualism, however, information about stakes is indirect. Information about stakes affects the context of use and can in turn play a role in fixing the content of the knowledge attributions .

So far then our results would seem to favor both contextualism and IRI equally. But this is not so. For ease of exposition, in what follows I assume a version of contextualism in which occurrences of the word 'knows' can express different relations (invoking varying epistemic standards). But what we say can also apply to the other versions of contextualism. I will also assume a traditional intellectualist contextualism. This is the view according to which each 'knowledge' relation is not sensitive to stakes. The criticisms I develop below are not applicable if we relax this assumption and adopt instead a version of contextualism where some 'knowledge' relations are sensitive to stakes.

Consider the results of study 2 (juxtaposed vignettes). In that study, subjects were given the pair of 'Typo' vignettes. They were then given a separate prompt where they were supposed to fill in the blanks. Let us focus on one version of the prompt (where the high stakes part goes before the low stakes part):

> (**) John [high stakes] needs to proofread his paper X times before he knows there are no typos, and Peter [low stakes] needs to proofread his paper Y times before he knows there are no typos.

people implicitly hold that knowledge is factive or that it entails belief, the explanation will certainly invoke the idea that knowledge is indeed factive and that it entails belief.

[66] My strategy somewhat resembles the following: (1) Intuitive knowledge attributions are sensitive to stakes (in a very broad sense) (2) This is better explained by IRI than contextualism. Therefore: IRI gains support. Jessica Brown pointed out to me that this strategy is similar to one of Stanley's (2005) strategies. My argument has a different structure, however, since I have not made use of (1). Instead, I collected experimental data and argued that the data supports the hypothesis that the folk behave as if IRI is true (this part might be more interesting to cognitive scientists than to some epistemologists). I then developed a further argument that the data is explained well by IRI and that it supports IRI over traditional contextualism.

Consistent with IRI, the X answers trended higher than the Y answers.[67] This bodes well for IRI, for these answers are the predicted ones on that view. But what about contextualism? The answer is not as clear.

To begin, we must appreciate that the prompt (★★) is a piece of writing that belongs to the context of the experiment and should be seen as a piece of the dialogue between the experimenter and the subject. It is not a speech made by any protagonist in the vignette. Notice now that if a subject writes different numbers for X and for Y, the two occurrences of 'knows' will express different relations reflecting the different epistemic standards at play. Thus, according to contextualism, we would have a single utterance of a sentence where different standards of knowledge are invoked. Although, there is not a general problem about context-sensitive words shifting content mid-discourse, I believe that the way this is happening here is somewhat difficult for the contextualist to explain. There are three reasons for this.

First, consider a common view among contextualists that what raises the standard for knowledge is salience of error (Cohen 2005; Lewis 1996; Schaffer 2004). If this is right, then for those subjects who wrote numbers X>Y in (★★) (most of them) error was salient in the first part of the sentence but it failed to be salient moments later in the second half of the sentence. This is hard to believe. It is not clear how the possibility of error can stop being salient all of a sudden in this manner. This is, at the very least, a very serious prima facie problem for the 'salience of error' account. Furthermore, if mention of high stakes in the first part of the sentence in the juxtaposed condition (★★) raises the epistemic standards, then we should expect the epistemic standards to rise and hence the answers to the second part to also rise. But this is not the case. If we compare the low stakes answer (second part, "Y") in condition (★★) (median=2) with the low stakes answer in the simple (non-juxtaposed) case from study 1 (median=2), we find that there is no statistically significant difference between those two figures.[68] This result puts some further pressure against the saliency account because it suggests that mention of high stakes does not raise the standards.

Here is the second reason why contextualists have a difficult time explaining our data (and this is independent of whether we accept the 'salience of error' account). Our data concerning (★★) suggest that, from the perspective of the contextualist, subjects can shift knowledge standards halfway through an utterance. And they will do this without taking back what they said at the beginning. After all, our subjects had plenty of time to change their answers to (★★). Now if it were possible for subjects to use 'knowledge' with different standards in a single utterance (without taking back the first use), then we should expect certain infelicitous utterances to be in fact felicitous. Suppose, for example, that A is in an epistemology class (high standards) while B is at a local bar (low standards) and both subjects have the same evidence available to them for the proposition that they have hands. If we can vary the standards within an

[67] Wilcoxon Signed Ranks Test (N=45) z=-4.68, p<.001, r=.7.
[68] Mann–Whitney (N=122) U=1533.000, z=-1.11, p=.266, r=–.1.

utterance in the way we see with (★★) we should be able to say things like 'A needs to ensure that he is not a brain in a vat before he knows that he has hands, but B simply needs to look with his eyes before he knows he has hands', or 'A does not know that he has hands but B does know he has hands', and even (in the first person): 'I didn't know that I had hands in epistemology class, but in the bar I knew that I had hands'. But these utterances are infelicitous, contrary to what we would expect if contextualism could really account for our data. These considerations then create problems for contextualists. They need to explain why knowledge standards could shift in (★★) (without subjects taking back the first part of the claim) but they could not in the cases we just saw.[69] And again, this is not a problem that arises for IRI.[70] So I think the data here favors IRI over contextualism.

Finally, recall the explanation given by people who wrote down different answers to the juxtaposed cases. Many of them said (roughly) that John should check more times than Peter because he has more at stake. This explanation is neatly explained by the idea that people implicitly accept RKP and that RKP is true. But if RKP is true, then knowledge is sensitive to stakes suggesting IRI. Contextualists need a different explanation. Contextualists usually want to explain the effect on stakes as a 'pre-semantic' phenomenon (DeRose 2004; Schaffer 2006). Thus, many contextualists would need another explanation for why subjects explain their answers in the way they do without appealing to the truth of RKP. How can the contextualist explain why the normative and knowledge responses were the same and how can they explain why only one person chose to use the word-form 'know' to explain their answers to the knowledge prompt? At this stage it looks like IRI has a better explanation of this data than contextualism.

In this section, I have suggested that contextualism will have some difficulties explaining the juxtaposed responses the subjects gave to (★★). I showed in several ways how the 'salience of error' mechanism will have some difficulty in accounting for the data. Also, independent of this mechanism, contextualists will have a difficult time explaining how standards can vary within a single utterance. Finally, they will have a difficult time explaining the written explanations subjects gave to (★★). In contrast, the data from answers to (★★) is explained in a straightforward way by IRI. I do not wish to say, however, that we have conclusive evidence against contextualism. Rather my claim is that IRI explains our data somewhat better than (traditional) contextualism. No doubt, further work needs to be carried out.

[69] Defenders of contextualism may explore the idea that the experiments involving the juxtaposed prompts are not ecologically valid.

[70] Stanley (2005) notes that contextualists about knowledge are committed to certain infelicitous sentences being felicitous. My strategy is similar except that the commitment of the contextualist arises from the sort of response they would presumably need to give to the specific data presented here. Thanks to Jessica Brown here.

9.7 Conclusion

In this chapter I have given some new evidence for the claim that ordinary attributions of knowledge are sensitive to practical interests in the sense of IRI (FS–IRI). I have done so mainly by introducing various manipulations on 'evidence-seeking' experiments. I have also argued that IRI provides an elegant explanation of the data and that the same cannot be said about traditional contextualist theories. The experiments presented here then are the first to date which, arguably, directly support IRI. Of course, I do not want to say that I have provided conclusive evidence for IRI, or against contextualism. Rather, I see the work here as adding some new considerations to that debate. On a broader note, I hope to have shown how through the new methods introduced here, we can gain insights into some central problems in epistemology.

Appendix

Typo-Low

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

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.

Ignorant-Typo-High

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 his paper. However, John is unaware of what is really at stake. He thinks the teacher does not care at all if there are some or even many typos in the paper. Although John would like to have no typos, he is unaware that it would be extremely bad for him if there is but a single typo in the paper.

PROMPT: How many times do you think Peter (John) has to proofread his paper before he knows that there are no typos? ____ times.
PROMPT (juxtaposed): Peter needs to proofread his paper ___ times before he knows there are no typos, and John needs to proofread his paper ___ times before he knows there are no typos.

Count-Low

Peter, a bright college student, has just won a contest sponsored by his bank. The prize is a chance to win some movie tickets. In order to win the tickets and as a publicity stunt, Peter is given a whole day to correctly count all the pennies in a jar located at his local branch. The jar contains around a hundred pennies. Peter wants the passes but it's not a big deal at all if he doesn't get the correct answer.

Count-High

Peter, a bright college student, has just won a contest sponsored by his bank. The prize is a chance to win a very large sum of money. In order to win the money and as a publicity stunt, Peter is given a whole day to correctly count all the pennies in a jar located at his local branch. The jar contains around a hundred pennies. It turns out there is a lot at stake for Peter. Peter is fully aware that if he answers correctly, he will win the prize and thereby be able to pay for his mother's life-saving operation (which they couldn't afford otherwise). So it's really important that Peter count the pennies correctly since his mother's life is at stake.

PROMPT ("knowledge"): Peter must count the pennies in the jar at least ___ times before he comes to know the correct answer.
PROMPT ("normative"): Peter should count the pennies in the jar at least ___ times before he declares his final answer to the judges.

Bibliography

Brown, J. 2008. "Subject Sensitive Invariantism and the Knowledge Norm for Practical Reasoning," *Noûs*, 42(2): 167–89.

—— forthcoming. "Experimental Philosophy, Contextualism and SSI," *Philosophy and Phenomenological Research*.

Buckwalter, W. forthcoming. "Knowledge Isn't Closed on Saturdays: A Study in Ordinary Language," *Review of Philosophy and Psychology*.

Chaiken, S. 1980. "Heuristic versus Systematic Information Processing and the Use of Source versus Message Cues in Persuasion," *Journal of Personality and Social Psychology*, 40: 111–20.

Cohen, S. 1986. "Knowledge and Context," *Journal of Philosophy*, 83: 574–83.

—— 2005. "Knowledge, Speaker, Subject," *Philosophical Quarterly*, 55(219): 199–212.

DeRose, K. 1992. "Contextualism and Knowledge Attributions," *Philosophy and Phenomenological Research*, 52(4): 913–29.

—— 2004. "The Problem with Subject Sensitive Invariantism," *Philosophy and Phenomenological Research*, 68(2): 346–50.

—— 2006. "Bamboozled by Our Own Words: Semantic Blindness and Some Arguments against Contextualism," *Philosophy and Phenomenological Research*, 73(2): 316–38.

—— 2009. *The Case for Contextualism: Knowledge, Skepticism and Context,* Vol. 1 (New York: Oxford University Press).

—— forthcoming. "Contextualism, Contrastivism and X-phi Surveys", *Philosophical Studies*.

Fantl, J. and McGrath, M. 2007. "On Pragmatic Encroachment in Epistemology," *Philosophy and Phenomenological Research*, 75: 558–89.

—— —— 2009. *Knowledge in an Uncertain World* (Oxford: Oxford University Press).

Feltz, A. and Zarpentine, C. forthcoming. "Do You Know More When it Matters Less?" *Philosophical Psychology*.

Frederick, S. 2005. "Cognitive Reflection and Decision Making," *Journal of Economic Perspectives*, 19: 25–42.

Gerken, M. (ms). "The roles of knowledge ascriptions in epistemic assessment."

Goldman, A. 2007. "Philosophical Intuitions: Their Target, Their Source and Their Epistemic Status," *Grazer Philosophische Studien*, 74: 1–26.

Hagafors, R. and Brehmer, B. 1983. "Does Having to Justify One's Judgments Change the Nature of the Judgment Process?" *Organizational Behavior and Human Performance*, 31: 223–32.

Hawthorne, J. 2004. *Knowledge and Lotteries* (New York: Oxford University Press).

—— and Stanley, J. 2008. "Knowledge and Action," *Journal of Philosophy*, 105(10): 571–90.

Heller, M. 1995. "The Simple Solution to the Problem of Generality," *Noûs*, 29(4): 501–15.

Kahn, B. and Baron, J. 1995. "An Exploratory Study of Choice Rules Favored for High Stakes Decisions," *Journal of Consumer Psychology*, 4(4): 305–28.

Kahneman, D. and Frederick, S. 2002. "Representativeness Revisited: Attribute Substitution in Intuitive Judgment," in T. Gilovich, D. Griffin and D. Kahneman (eds.), *Heuristics and Biases: The Psychology of Intuitive Judgment* (New York: Cambridge University Press).

—— and Tversky, A. 1996. "On the Reality of Cognitive Illusions," *Psychological Review*, 103 (3): 582–91.

Kassin, S. M., Castillo, S. R. and Rigby, S. 1991. "The Accuracy–Confidence Correlation in Eyewitness Testimony: Limits and Extensions of the Retrospective Self-Awareness Effect," *Journal of Personality and Social Psychology*, 5: 698–707.

Knobe, J. 2010. "Person as Scientist, Person as Moralist," *Behavioral and Brain Sciences*, 33: 315–39.

Lewis, D. 1979. "Scorekeeping in a Language Game," *Journal of Philosophical Logic*, 8, 339–59.

—— 1996. "Elusive Knowledge," *Australasian Journal of Philosophy*, 74(4): 549–67.

Luce, M. F., Bettman, J. R., and Payne, J. W. 1997. "Choice Processing in Emotionally Difficult Decisions," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23: 384–405.

Ludlow, P. 2005. "Contextualism and the New Linguistic Turn in Epistemology," in G. Preyer and G. Peter (eds.), *Contextualism in Philosophy: Knowledge, Meaning and Truth* (New York: Oxford University Press).

McAllister, D.W., Mitchell, T. R. and Beach, L.R. 1979. "The Contingency Model for Selection of Decision Strategies: An Empirical Test of the Effects of Significance, Accountability, and Reversibility," *Organizational Behavior and Human Performance*, 24(2): 228–44.

May, J. W., Sinnott-Armstrong, J., Hull, G. and Zimmerman, A. forthcoming. "Practical Interests, Relevant Alternatives and Knowledge Attributions," *Review of Philosophy and Psychology*.

Mayseles, Q. and Kruglanski, A. W. 1987. "What Makes You So Sure? Effects of Epistemic Motivation on Judgmental Confidence," *Organizational Behavior and Human Decision Processes*, 39: 162–83.

Moore, G. E. 1939. "Proof of an External World," *Proceedings of the British Academy*, 25: 273–300.

Nagel, J. 2008. "Knowledge Ascriptions and the Psychological Consequences of Changing Stakes," *Australasian Journal of Philosophy*, 86: 279–94.

—— forthcoming. "Epistemic Anxiety and Adaptive Invariantism," *Philosophical Perspectives*.

Neta, R. 2002. "S knows that P," *Noûs*, 36(4): 663–81.

—— 2007. "Anti-Intellectualism and the Knowledge–Action Principle," *Philosophy and Phenomenological Research*, 75: 1.

Pinillos, N. A. forthcoming. "Recent Work in Experimental Epistemology," *Philosophy Compass*.

—— Smith, N., Nair, G., Marchetto, P. and Mun, C. 2011. "Philosophy's New Challenge: Experiments and Intentional Action," *Mind and Language*, 26(1): 115–39.

Rieber, S. 1998. "Skepticism and Contrastive Explanation," *Noûs*, 32(2): 189–204.

Rozelle, R. M. and Baxter, J. C. 1981. "Impression formation under conditions of accountability and self salience," ms, University of Houston.

Rysiew, P. 2007. "Speaking of Knowing," *Noûs*, 41(4) 627–62.

Schaffer, J. 2004. "From Contextualism to Contrastivism," *Philosophical Studies*, 119(1–2): 73–103.

—— 2006. "The Irrelevance of the Subject: Against Subject Sensitive Invariantism," *Philosophical Studies*, 127: 87–107.

—— and Knobe, J. forthcoming. "Contrastivism Surveyed," *Noûs*.

Shatz, M., Wellman, H. M. and Silber, S.1983. "The Acquisition of Mental Verbs: A Systematic Investigation of the Child's First Reference to Mental State," *Cognition*, 14: 301–21.

Slovic, P., Finucane, M., Peters, E. and McGregor, D. 2001. "The Affect Heuristic," in T. Gilovich, D. Griffin, and D. Kahneman (eds.), *Intuitive Judgment: Heuristics and Biases* (New York: Cambridge University Press).

Stanley, J. 2005. *Knowledge and Practical Interests* (New York: Oxford University Press).

Szabó Gendler, T. 2008. "Alief and Belief," *Journal of Philosophy*, 105(10): 634–63.

—— and Hawthorne, J. 2005. "The Real Guide to Fake Barns: A Catalogue of Gifts for Your Epistemic Enemies," *Philosophical Studies*, 124: 331–52.

Tetlock, P. E. 1992. "The Impact of Accountability on Judgment and Choice: Toward a Social Contingency Model," *Advances in Experimental Social Psychology*, 25: 331–76.

Weatherson, B. (2011). "Defending Interest-Relative Invariantism", *Logos & Episteme* 2(4): 591–609.

Williamson, T. 2000. *Knowledge and Its Limits* (New York: Oxford University Press).

Zimmerman, D. 2007. "The Privileged Present: Defending an 'A-Theory' of Time," in T. Sider, J. Hawthorne and D. Zimmerman (eds.), *Contemporary Debates in Metaphysics* (Oxford: Blackwell).