

The Instability of Philosophical Intuitions: Running Hot and Cold on Truetemp

Stacey Swain, Joshua Alexander, and Jonathan M. Weinberg¹

Indiana University, Bloomington

(Forthcoming in *Philosophy and Phenomenological Research*)

Abstract:

A growing body of empirical literature challenges philosophers' reliance on intuitions as evidence based on the fact that intuitions vary according to factors such as cultural and educational background, and socio-economic status. Our research extends this challenge, investigating Lehrer's appeal to the Truetemp Case as evidence against reliabilism. We found that intuitions in response to this case vary according to whether, and which, other thought experiments are considered first. Our results show that compared to subjects who receive the Truetemp Case first, subjects first presented with a clear case of knowledge are less willing to attribute knowledge in the Truetemp Case, and subjects first presented with a clear case of non-knowledge are more willing to attribute knowledge in the Truetemp Case. We contend that this instability undermines the supposed evidential status of these intuitions, such that philosophers who deal in intuitions can no longer rest comfortably in their armchairs.

¹ Each author contributed fully to this project. We would like to thank our colleagues from the Experimental Epistemology Laboratory (EEL) at Indiana University for ongoing discussion of our work: Cameron Buckner, Stephen Crowley, Chad Gonnerman, and David Landy. We would also like to thank John Alexander, Melinda Fagan, Adam Feltz, Mark Kaplan, Adam Leite, Thomas Nadelhoffer, Eddy Nahmias, Philip Robbins, and an anonymous referee for this journal, as well as participants at the 98th Annual Meeting of the Southern Society for Philosophy and Psychology, the 32nd Annual Meeting of the Society for Philosophy and Psychology, the first On-line Philosophy Conference, and the 3rd Annual EEL Mini-Conference for valuable feedback on earlier drafts of this paper.

Introduction:

A growing body of empirical literature challenges philosophers' reliance on intuitions as evidence based on the fact that intuitions vary according to philosophically irrelevant factors, such as cultural and educational background or affective state.² Our research extends this challenge, demonstrating that intuitions vary according to whether, and which, other thought experiments are considered first. We critique the use of intuitions, such as those generated by Lehrer's Truetemp Case, as evidence, on the grounds that intuitions about this case are easily manipulated: compared to subjects who receive the Truetemp Case first, subjects first presented with a clear case of knowledge are less willing to attribute knowledge in the Truetemp Case, and subjects first presented with a clear case of non-knowledge are more willing to attribute knowledge in the Truetemp Case.

1. Intuitions as evidence: The Truetemp Case

Epistemological reliabilism typically holds that a person's true belief that *p* counts as knowledge just in case it is caused, or causally sustained, by a reliable cognitive process.³ Keith Lehrer's Truetemp Case is standardly appealed to as an argument against reliabilism:

² See Jonathan M. Weinberg, Shaun Nichols, and Stephen Stich, "Normativity and Epistemic Intuitions", *Philosophical Topics* Vol. 29 (2001): pp. 429-460; Shaun Nichols, Stephen Stich, and Jonathan M. Weinberg, "Metaskepticism: Meditations in Ethno-Epistemology", in Stephen. Luper, (ed.), *The Sceptics*, (Ashgate Publishing, 2003): pp. 227-247; Edouard Machery, Ron Mallon, Shaun Nichols, and Stephen Stich, "Semantics, Cross-cultural Style", *Cognition*, Vol. 92, No. 3, pp. B1-B12; and Shaun Nichols and Joshua Knobe, "Moral Responsibility and Determinism: The Cognitive Science of Folk Intuitions", forthcoming in *Nous*.

³ Reliabilism traces its roots to Frank Ramsey, "Knowledge", *Foundations of Mathematics and Other Logical Essays*, (Routledge, 1931). See also, A. Phillips Griffiths, *Knowledge and Belief*, (Oxford University Press, 1967); Alvin Goldman, *Epistemology and Cognition*, (Harvard University Press, 1986); Alvin Goldman, "Reliabilism", in Jonathan Dancy and Ernest Sosa, (eds.), *A Companion to Epistemology*, (Blackwell, 1992): pp. 433-436; Alvin Goldman, "Naturalistic Epistemology and Reliabilism", in Peter French *et al.* (eds.), *Midwest Studies in Philosophy* Vol. 19, (University of Minnesota Press, 1994): pp. 301-320. This is a very general description of reliabilism. There are many, and varying, versions of reliabilism; however, here, and in what follows, we gloss over such distinct versions of the position. Since our purpose is to critique the use of intuitions, and not any particular version of reliabilism, the general description suffices. Additionally, in what follows, we gloss over the distinction between

Suppose a person, whom we shall name Mr. Truetemp, undergoes brain surgery by an experimental surgeon who invents a small device which is both a very accurate thermometer and a computational device capable of generating thoughts. The device, call it a tempucomp, is implanted in Truetemp's head so that the very tip of the device, no larger than the head of a pin, sits unnoticed on his scalp and acts as a sensor to transmit information about the temperature to the computational system of his brain. This device, in turn, sends a message to his brain causing him to think of the temperature recorded by the external sensor. Assume that the tempucomp is very reliable, and so his thoughts are correct temperature thoughts. All told, this is a reliable belief-forming process. Now imagine, finally, that he has no idea that the tempucomp has been inserted in his brain, is only slightly puzzled about why he thinks so obsessively about the temperature, but never checks a thermometer to determine whether these thoughts about the temperature are correct. He accepts them unreflectively, another effect of the tempucomp. Thus, he thinks and accepts that the temperature is 104 degrees. It is. Does he know that it is?⁴

According to reliabilism, if a person's true belief that *p* is caused by a reliable cognitive process, then that belief qualifies as knowledge. Mr. Truetemp's temperature beliefs are caused by a reliable cognitive process. Therefore, according to reliabilism, Mr. Truetemp does know it is 104 degrees. But Lehrer claims that there is something lacking in Mr. Truetemp's epistemic position, such that his temperature beliefs do not count as knowledge. Purportedly, if we consider this case, we will have the intuition that Mr. Truetemp does not know that it is 104 degrees. Reliabilism's inability to account for this intuition is supposed to be reason to reject reliabilism.

Generally, philosophers accept appeals to intuitions about the Truetemp Case as evidence against reliabilism. Even Alvin Goldman grants that Lehrer's Truetemp Case presents a serious objection to reliabilism.⁵ According to standard practice, a philosophical claim is *prima facie* good to the extent that it accords with our intuitions, *prima facie* bad to the extent that it does

a belief's being causally produced and a belief's being causally sustained by a cognitive process because the distinction does not matter for our purposes.

⁴ Keith Lehrer, *Theory of Knowledge*, (Westview Press, 1990): pp. 163-164.

⁵ See Alvin Goldman, "Reliabilism", *op. cit.* and "Naturalistic Epistemology and Reliabilism", *op cit.*

not. Given that intuitions about thought-experiments are standardly taken as reasons to accept or reject philosophical theories, then we should be interested in finding out what the relevant intuitions are.

Experimental philosophers have begun conducting empirical research to find out what intuitions are generated in response to certain cases. But rather than supporting and explaining the practice of appealing to intuitions as evidence, the results of this research challenge the legitimacy of appealing to intuitions. Weinberg, Nichols, and Stich revealed that epistemological intuitions vary according to factors such as cultural and educational background; Machery *et al.* document a similar cultural variation in semantic intuitions; and Nichols and Knobe have discovered that the affective content of a thought-experiment can influence whether subjects have compatibilist or incompatibilist intuitions.

To the extent that intuitions are sensitive to these sorts of variables, they are ill-suited to do the work philosophers ask of them. Intuitions track more than just the philosophically-relevant content of the thought-experiments; they track factors that are irrelevant to the issues the thought-experiments attempt to address. The particular socio-economic status and cultural background of a person who considers a thought experiment should be irrelevant to whether or not that thought-experiment presents a case of knowledge. Such sensitivity to irrelevant factors undermines intuitions' status as evidence. Since the intuitions generated by the Truetemp Case are sensitive to whether, and which, other thought-experiments are considered first, these intuitions are susceptible to manipulation. Evidence so unstable risks being discounted as not truly evidence at all. Furthermore, given that intuitions vary in these ways, there is unlikely to be

a fixed set of intuitions about a particular thought-experiment to which we can appeal.⁶ Finally, even if one were to grant that, in principle, intuitions can be used as evidence in philosophy, these results suggest that, at this time, we cannot tell *which* intuitions can safely be deployed.

2. Empirical Results

2.1. Method

220 students attending undergraduate classes at a large, Midwestern university were recruited for the study (136 men and 83 women).⁷ Subjects were randomly assigned to eight different survey versions.⁸ The eight survey versions differed only in the order in which the four thought-experiments were presented.

Each of the four thought-experiments presents a story about how a person comes to have a true belief, and asks subjects to respond to a statement attributing knowledge to the person in the

⁶ This possibility has been raised by Jonathan M. Weinberg, Shaun Nichols and Stephen Stich and by Robert Cummins. See, Jonathan M. Weinberg, Shaun Nichols, and Stephen Stich, “Normativity and Epistemic Intuitions”, *op. cit.*; and Robert Cummins, “Reflection on Reflective Equilibrium”, in Michael R. DePaul and William Ramsey, (eds.), *Rethinking Intuition*, (Rowman and Littlefield, 1998): pp. 113-129.

⁷ 228 students completed the survey; 8 were excluded due to responses to the screening question that suggested that they were working with a different conception of ‘knows’ than the one of interest to philosophers. (See below for discussion of the Coinflip Case as a screening question.) 1 student did not report gender. Participants averaged 20.6 years old. Additionally, participants averaged 3.8 semesters of college. The number of students assigned to each survey version ranged from 26 to 31. The survey consisted of two pages. The first page included basic instructions and four thought experiments. Students were asked to indicate, for each thought-experiment, to what extent they agreed or disagreed with a target statement attributing knowledge to the protagonist of the thought-experiment. The second page collected basic demographic data including age, gender, semesters of college completed, and the number of college courses taken in philosophy, mathematics, physical sciences, social sciences, and art/literature. Researchers visited 12 undergraduate classes to recruit students for the survey. At the end of each class, the researcher(s) addressed the students, explaining the survey and inviting them to participate. Students were given randomly assigned versions of the survey. Students who chose to participate turned in their completed surveys as they left the room. Students’ willingness to agree to a statement attributing knowledge to the subjects in each of the four thought experiments was coded such that higher numbers correspond to greater willingness to attribute knowledge (1: strongly disagree, 2: disagree, 3: neutral, 4: agree, 5: strongly agree). The data was then analyzed using SPSS 13.0.

⁸ One version of the actual survey is presented in Appendix 1.

story using a 5-point Likert Scale, a standard measure of attitude toward a proposition.⁹ The central thought-experiment in the survey involves a man named Charles, modeled after Lehrer's Mr. Truetemp.¹⁰ The three other thought-experiments, which we will briefly discuss here, include a clear case of non-knowledge, a clear case of knowledge, and a Ginet-style fake barn case.

The clear case of non-knowledge involves Dave, a man who believes he can predict when a coin will land heads.¹¹ Dave's predictions of the outcome of coinflips are no better than chance, but in this case he happens to get it right. This thought-experiment was included for two reasons. First, willingness to attribute knowledge to Dave suggests that subjects are taking 'knows' to mean 'has a strong sense of subjective certainty.'¹² On this basis, 8 subjects were excluded from the analysis because they are presumed to be working with a meaning of 'knows' other than the meaning that interests epistemologists. (Their exclusion did not significantly affect the results of the analysis.¹³) Second, we antecedently expected that Dave has a true belief that does not count as knowledge. This expectation was confirmed: only 8 of the 228 subjects were willing to attribute knowledge in that case. The main experimental hypothesis was that subjects' responses to the Truetemp Case would vary significantly based on which other cases were presented before it. A pilot study indicated that responses to this Truetemp Case are mixed; some subjects are willing to attribute knowledge to Charles, some are not. Since the Coinflip Case obviously did not involve knowledge, it was expected to make the Truetemp Case seem more plausibly a case

⁹ See Appendix 1.

¹⁰ The full case is presented in Appendix 1.

¹¹ We borrow this case from Johnathan M. Weinberg, Shaun Nichols, and Stephen Stich, "Normativity and Epistemic Intuitions", *op. cit.* The full case is presented in Appendix 1.

¹² See Johnathan M. Weinberg, Shaun Nichols, and Stephen Stich, "Normativity and Epistemic Intuitions", *op. cit.*

¹³ See footnote 17 for one small exception.

of knowledge by comparison. Compared to subjects who receive the Truetemp Case first, subjects who are first presented with a clear case of non-knowledge were expected to be more willing to attribute knowledge in the Truetemp Case.

The rationale for including the clear case of knowledge, involving a chemist named Karen, was similar.¹⁴ It was expected that a very clear case of knowledge would make the Truetemp Case seem less plausibly a case of knowledge by comparison. The Chemist Case is a clear case of non-controversial knowledge of a trained expert gaining information by reading a leading scientific journal.

The Ginet-style fake barn case involves a character Suzy who is driving through a countryside populated with fake barns. This case was expected to generate mixed intuitions; with some subjects willing to attribute knowledge, and others not. Since the other cases were designed to test the effects of presenting a clear case of knowledge and a clear case of non-knowledge before the Truetemp Case, we included the last case to test the effects of presenting a mixed case before the Truetemp Case.

2.2. Main Findings

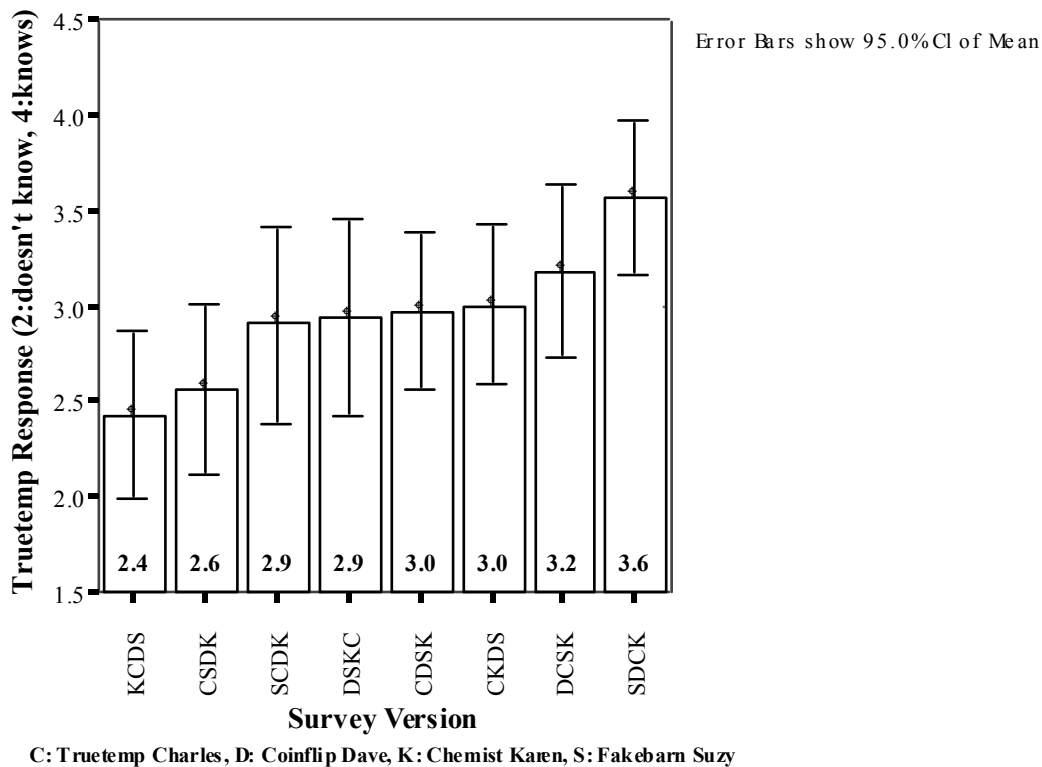
The main hypothesis driving this study was that subjects' willingness to attribute knowledge to Charles in the Truetemp Case will vary depending on whether, and which, other cases are presented before it. More specifically, we had two sub-hypotheses: compared to subjects who receive the Truetemp Case first, 1) subjects who are first presented with a clear case of

¹⁴ The full case is presented in Appendix 1.

knowledge (the Chemist Case) are less willing to attribute knowledge in the Truetemp Case, and
 2) subjects who are first presented with a clear case of non-knowledge (the Coinflip Case) are
 more willing to attribute knowledge in the Truetemp Case.

Mean Truetemp Response by Survey Version

ANOVA $p=.012$

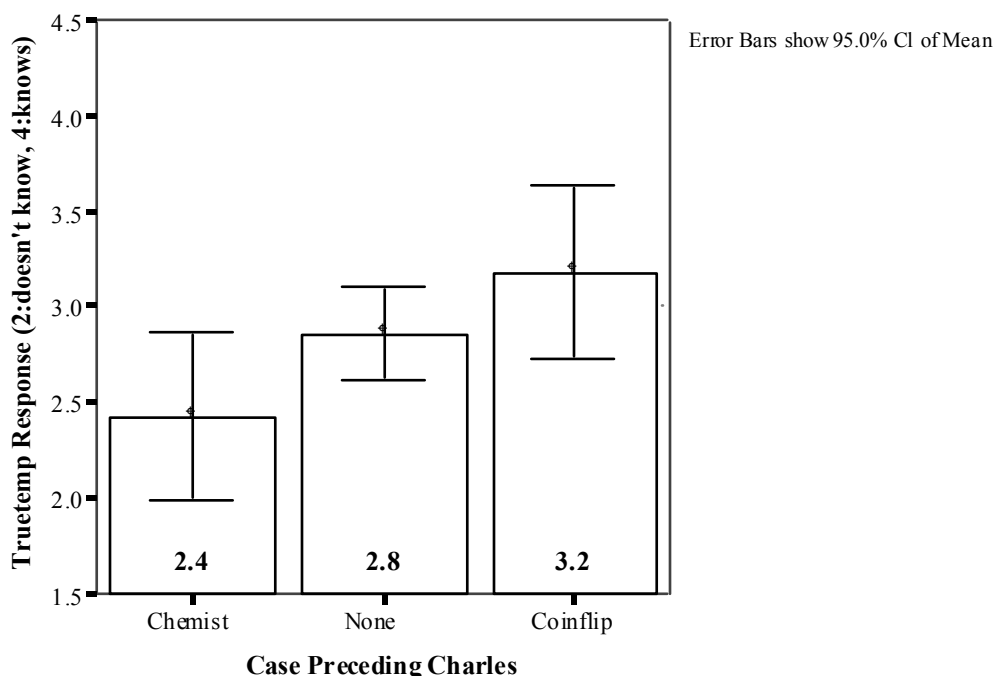


Our hypotheses were confirmed in the analysis of the data. Analysis of variance based on survey version in the Truetemp Case revealed a significant effect ($p = .012$).¹⁵ Subjects' willingness to attribute knowledge in the Truetemp Case varied significantly across survey versions.¹⁶

¹⁵ All of the results presented are based on analysis in which 8 subjects who were willing to attribute knowledge in the Coinflip Case were excluded. Including those cases in the analysis did not significantly affect the results.

Mean Truetemp Response by Case Preceding Charles

ANOVA $p=.048$



X-axis values represent survey versions in which the Truetemp Case is presented first, presented second following the Chemist Case, and presented second following the Coinflip Case.

Additionally, we confirmed that subjects who are first presented with a clear case of non-knowledge are significantly more willing to attribute knowledge in the Truetemp Case ($p = .043$).¹⁷ This is an important finding because the increase from 2.64 to 3.31 crosses the threshold

¹⁶ The survey versions varied according to the order in which the cases were presented, and were coded as follows, where C: Truetemp Charles, D: Coinflip Dave, K: Chemist Karen, and S: Fakebarn Suzy, 1: CDSK, 2: DCSK, 3: SDCK, 4: DSKC, 5: SCDK, 6: KCDS, 7: CSDK, 8: CKDS.

¹⁷ This is the result of a t-test (including the subjects who attributed knowledge to Coinflip Dave) comparing a survey version where Charles is first—CSDK (mean response to Charles: 2.64) to survey version where Charles directly follows Dave—DCSK (mean response to Charles: 3.31). The same t-test was performed excluding the subjects who attributed knowledge to Charles and the result was similar ($p = .057$). T-tests on the other versions where Charles appeared first or immediately following the Coinflip Case showed the same trend, although results were marginally significant or not significant. A t-test comparing the composite of the three versions in which

of neutrality (3.0 = neutral). Subjects who encounter Charles before any other cases are unwilling to attribute knowledge to him; however, subjects who are first presented with the Coinflip Case will attribute knowledge to Charles. Intuitions about the Truetemp Case reverse direction depending on whether the case is presented after a case of clear non-knowledge. The fact that people's intuitions about particular thought-experiments vary based on what other things they have been thinking about recently is troubling. Philosophers who rely on thought-experiments should be especially concerned about findings that indicate that, at least in some cases, subjects' intuitions are easily influenced.

Additionally, as predicted, subjects who are first presented with a clear case of knowledge¹⁸ are significantly less willing to attribute knowledge in the Truetemp Case (from 3.00 to 2.41, $p = .054$).¹⁹ Interestingly, this is exactly what happens in Lehrer's *A Theory of Knowledge*, where the Truetemp Case is first introduced. In the section immediately preceding presentation of the Truetemp Case, Lehrer discusses paradigm cases of knowledge: perceptual knowledge, knowledge arrived at through communication with others, and knowledge of mathematics.²⁰ We are not suggesting that Lehrer *intentionally* manipulated any evidence appealed to as part of his case against epistemological reliabilism; rather, we are concerned that philosophers might be manipulating their own results *without even being aware that such manipulation is taking place*.

Charles appears first to the version where Charles directly follows Coinflip Dave was not significant ($p = .20$); however, the same t-test was performed including the 4 subjects who were willing to attribute knowledge to Coinflip Dave (1 subject with the Charles first version, 3 subjects with the Dave first version), and the result was marginally significant ($p = .084$).

¹⁸ Mean response to the Chemist Case = 3.94.

¹⁹ This is the result of a t-test comparing survey version where Charles follows directly Karen—KCDS (mean response to Charles: 2.41) to survey version in which Charles is first—CKDS (mean response to Charles: 3.00). T-tests on the other versions where Charles appeared first or immediately following the Chemist Case showed the same trend, although results were marginally significant or not significant. A t-test comparing the composite of the three versions in which Charles appears first to the version where Charles directly follows Karen was marginally significant ($p = .079$).

²⁰ Keith Lehrer, *Theory of Knowledge*, op. cit, pp. 159-161.

Our findings suggest that Lehrer's readers' unwillingness to attribute knowledge to Mr. Truetemp may be influenced by the preceding cases; if the Truetemp Case were presented without those preceding cases, readers might be less confident about denying that Mr. Truetemp has knowledge. If one is to take the Truetemp Case as evidence against reliabilism, one must find some way of explaining this anomaly.

2.3. Additional Findings

The purpose of including the Fakebarn Case was somewhat exploratory; we suspected that this would be a case in which subjects' willingness to attribute knowledge would be mixed. We were interested to find out what subjects' intuitions about the case *actually* are, and we wondered whether intuitions about the Fakebarn Case would vary, similar to the Truetemp Case, based on which cases were presented before it. The mean response to the Fakebarn Case was 3.6 (compared to 1.6 for the Coinflip Case, and 3.9 for the Chemist Case).²¹

We were not sure whether the Fakebarn Case would demonstrate the same flexibility as the Truetemp Case with respect to which cases are presented before it. We found that subjects' intuitions about this case were, given the Truetemp Case's lability, surprisingly stable across presentation position.²² This raises an interesting question for the philosopher who relies on

²¹ Our colleague, Adam Leite, pointed out that it is difficult to draw any philosophical conclusions based solely on the mean response to our version of the Fakebarn Suzy case. This is because our version does not include the counterfactual claim that had Suzy been looking at a Fakebarn, she still would have believed that it was a real barn. Inclusion of that claim may affect subjects' willingness to attribute knowledge to Suzy. This point highlights the potential for experimental philosophy to identify which aspects of thought experiments are most salient.

²² Across the eight survey versions, where the Fakebarn Case appears in each position, mean responses varied only from 3.5 to 3.8. The ANOVA on response to the Fakebarn Case based on survey version did not reveal a significant effect. In fact, it suggests that responses to the Fakebarn Case are immutable with respect to whether, and which, other cases precede it (ANOVA, $p = .980$). A t-test comparing the survey version that generated the highest

intuitions: which intuitions, if any, are resistant to the potential effects of irrelevant factors? It is worth pointing out that any attempt to answer this question will need to rely on empirical research about intuitions. There is no *a priori* way to figure which intuitions are stable.

We were also interested in whether responses to any of the cases vary according to gender, age, semesters of college completed, or number of courses taken in any particular area. Nothing of statistical significance was found other than a statistically significant gender effect on willingness to attribute knowledge in the Coinflip Case (t-test, $p = .007$). Males had a slightly lower mean response to the case than females (1.53 versus 1.83), although almost all men and almost all women were unwilling to attribute knowledge. This difference appears simply to reflect the male subjects' greater willingness to use the whole range of the Likert scale in general; we do not think that this reflects any meaningful difference in intuitions. This suggests that a greater percentage of women than men were working with the "subjective certainty" meaning of 'knows'. The lack of evidence of additional effects may be due to the relative homogeneity of our sample with regard to these demographics.

3. Objections and Replies

In this section we consider two types of objections to our argument against the use of intuitions as evidence in epistemology. The first two objections challenge the representativeness of our findings and the scope of our conclusion. The third objection challenges the very way we interpret our data about the Truetemp Case.

response to the Fakebarn Case with the version that generated the lowest response to the Fakebarn Case was also insignificant ($p=.385$). It is worth noting that we did not have a survey version in which the Fakebarn Case followed only the Chemist Case. Thus, our findings of stability are somewhat hedged as there remains a possibility that responses to the Fakebarn Case might be different in this particular case.

Objection 1: These results are reason to reject intuitions about the Truetemp Case as evidence against reliabilism. However, the instability of some intuitions under some circumstances is not reason to reject intuitions as evidence more generally. After all, even some perceptual judgments are unreliable under some circumstances, but this is not reason to reject perception as a source of evidence. Ernest Sosa offers the analogy to perception in response to an earlier version of this paper.²³ Sosa writes:

One would think that the ways of preserving the epistemic importance of perception in the face of such effects on perceptual judgments would be analogously available for the preservation of the epistemic importance of intuition in the face of such effects on intuitive judgments. The upshot is that we have to be *careful* in how we use intuition, not that intuition is useless.²⁴

Reply: It is not clear what it is about the Truetemp Case that makes it sensitive to whether, and which, other thought experiments are presented first. If it were clear under which circumstances intuitions are unstable, we might be able to avoid relying on those intuitions under those circumstances. Thus, the real problem is that we simply do not know when intuitions are susceptible to these effects and when they are not. This is where the analogy to perception fails. We are aware of the great majority of the circumstances under which perceptual judgments are likely to be unreliable. For instance, we know that visual perception requires a certain amount of illumination, and visual perception itself provides us with knowledge of whether enough illumination is present. A dimly lit room does not appear gray; it appears dimly lit. Thus, the fact that we cannot discern whether the cat is in the room when the light is off does not justify skepticism about our ability to discern whether the cat is in the room when the light is on. At this time, we don't know what is the parallel for intuition of making sure that the light is on; that is, we do not know which are the circumstances that render intuition reliable or unreliable.

²³ Ernest Sosa, "Experimental Philosophy and Philosophical Intuition", *Philosophical Studies*, (forthcoming).

²⁴ Ernest Sosa, "Experimental Philosophy and Philosophical Intuition", *op. cit.*, mss. pp. 9.

With perception, by contrast, we are aware of the dimensions of variance and how to compensate for them. We know to turn our heads toward the speaker if we cannot hear well, or to squint if we are trying to read a distant road sign, or to cleanse the palate before evaluating a fine wine. What our research indicates is that we do not have analogous knowledge relating to our practice of relying on intuitions.

It is a legitimate, empirically-testable hypothesis that there is something about the Truetemp case that makes it unrepresentative of intuitions in general. However, as increasing amounts of empirical data raise trouble along different dimensions for intuitions about thought-experiments of various sorts, it becomes less reasonable to suppose that the intuitions experimental philosophers attack are an unrepresentative lot. Similarly, it becomes less reasonable to suppose that these are the only circumstances under which intuitions are unstable. As the empirical case against particular intuitions expands, it falls on those philosophers who wish to continue to employ intuitions as evidence to demonstrate that intuitions about their favorite thought-experiments are not susceptible to this, and other, problematic effects. Sosa may be right that the upshot of the instability of intuition is that we ought to be careful; the problem is that it is not clear what it means to be careful when it comes to intuition. Thus, it falls on those philosophers who wish to continue relying on intuitions to figure out what it means to be careful. We do not rule out *a priori* the possibility that they will be able to do so. What we rule out is the possibility that they will be able to do so *a priori*.

Objection 2: These results show that the intuitions of college students are susceptible to a problematic order effect. However, the philosophical practice of relying on intuitions as

evidence is concerned with the intuitions of professional philosophers. Unlike the intuitions of college students, the considered judgments of philosophers may not be responsive to features irrelevant to the issues that thought-experiments are designed to address. That is, they may not be susceptible to this order effect. Therefore, these results don't seriously threaten the practice of relying on intuitions (here, philosophers' intuitions) as evidence in philosophical practice.

Reply: First, it is important to note that this objection is not one that is available to all philosophers who employ intuitions as evidence in philosophical argumentation. Frank Jackson, for example, claims to actually poll his students to get evidence about their intuitions about philosophical cases, clearly suggesting that he is not taking there to be a relevant distinction between professional philosophers and non-philosophers. Second, while it is the case that our results don't *directly* confront the practice of appealing to philosophers' intuitions as evidence, that doesn't mean that they don't confront the practice at all. We take ourselves to be in the position to indirectly confront the practice of appealing to philosophers' intuitions as evidence on the basis of an inference from the demonstrated instability of the intuitions of non-philosophers to a parallel instability of philosophers' intuitions. Now, it might be objected that, in the absence of empirical support for the claim that there is no difference between the intuitions of philosophers and non-philosophers, such an inference is suspect. However, we think that this confuses the dialectical burden. The burden of proof is on the person who would claim that there is a substantive difference between philosophers' intuitions and those of non-philosophers. In the absence of empirically-defensible evidence that philosophers' intuitions work significantly differently from those of non-philosophers, the inference is warranted. To claim the opposite—that is, to simply presume some difference—is not to go very far towards responding to the

challenge that our results pose to the practice of relying on intuitions as evidence. Sure, it's *possible* that philosophers' intuitions are different. But, until philosophers provide significant empirical evidence that they are, this response is insubstantial.²⁵

Sometimes it is objected that, even in the absence of empirical evidence, we have *prima facie* reason to think that philosophers' intuitions will not display this kind of instability: philosophers are more careful and attentive in the reflection that they give to thought-experiments than are non-philosophers.²⁶ But, even if we grant that philosophers' reflection is careful and attentive in a way that non-philosophers' reflection is not, it is unclear how such reflection would render philosophers' intuitions immune to this kind of instability. Our evidence suggests that reflection, itself, will be subject to these same factors. Reflection must start somewhere, after all, and our evidence suggests that those starting points are sensitive to things that philosophers have not previously expected them to be sensitive to. For, after all, we would predict that what other cases come up in during reflection, and in what order they come up, will affect the intuitions of the later cases. The proponent of this line of objection needs reflection to eliminate our effects—but reflection is at least as likely to partake of those effects without reducing them. And it is worth noting that this claim—that reflection is somehow a path-independent psychological

²⁵ There seems to be another way of conceiving of the philosophical practice of relying on intuitions as evidence. One might suggest that when an individual philosopher relies on intuitions as evidence, she is merely relying on her own personal intuitions. There are a number of problems with this suggestion. First, there is little evidence to suggest that individual philosophers actually take themselves to be appealing only to their own intuitions. Second, an argument that relied solely on appeal to the intuitions of the author of that argument would be quite weak. Unless the author is supposing that the relevant intuitions are widely shared, then there is little hope of convincing anyone. If one's readers reject—or fail to share—one's premises, then one's arguments will be unconvincing. However, if the author *is* supposing that the relevant intuitions are widely shared, then the author should be understood to be appealing to the intuitions of all philosophers or to non-philosophers (or both). Thus, this reply either collapses into the one mentioned above, in which case it is open to the same reply, or it is open more directly to the challenge presented by our empirical study.

²⁶ This objection was raised by an anonymous referee for this journal.

process, that is itself immune from primacy and context effects—is itself a significant empirical claim of the sort whose truth cannot be ascertained from the armchair.

Objection 3: The argument that these findings undermine epistemologists' reliance on intuitions rests on the assumption that the factors according to which intuitions vary (such as whether, and which, other thought-experiments are considered first) really are irrelevant to the thought-experiments from the epistemologists' point of view. But there is an extant epistemological view that easily accommodates these findings: contextualism. According to contextualism,

The truth conditions of knowledge ascribing and knowledge denying sentences (sentences of the form 'S knows that P' and 'S doesn't know that P' and related variants of such sentences) vary in certain ways according to the contexts in which they are uttered. What so varies is the epistemic standards that S must meet (or, in the case of a denial of knowledge, fail to meet) in order for such a statement to be true.²⁷

It is relatively easy to devise a form of contextualism that is consistent with these findings, depending on how the 'relevant context in which [the sentences] are uttered' is understood. In normal discourse, we typically presuppose that questions are at least minimally worth asking. That is, when someone asks 'Does Karen know that P?' it must be in some sense possible that she does not. Contextualists typically endorse some form of what David Lewis called the *rule of accommodation*.²⁸ According to the rule of accommodation, subjects will adjust their epistemic standards to accommodate those of their interlocutors. Thus, the fact that the Chemist Case is such an obvious candidate for knowledge attribution will cause subjects to raise the standards for attributing knowledge. Similarly, because the Coinflip Case is such an obvious candidate for withholding knowledge attribution, subjects will lower the standards for attributing knowledge in

²⁷ Keith DeRose, "Solving the Skeptical Problem" reprinted in K. DeRose and Ted A. Warfield, (eds.), *Skepticism: A Contemporary Reader*, (Oxford University Press: 1999): p. 187.

²⁸ David Lewis, "Scorekeeping in a Language Game", *Journal of Philosophical Logic* Vol. 8 (1979): pp.339-59; and David Lewis, "Elusive Knowledge", in E. Sosa and J. Kim, (eds.), *Epistemology: An Anthology*, (Blackwell: 2000): pp. 503-516.

response to the presupposition that the question is worth asking. Because contextualism claims that the types of cases being considered alter the context in which the standards of knowledge attribution are assessed, contextualism would seem to predict the empirical results reported here. Thus, the supposed instability of intuitions about this thought-experiment is not reason to discontinue relying on them as evidence; rather it is properly understood as corroborating evidence for contextualism.

Reply: The truth or falsity of epistemic contextualism is not of concern to us here. We are not objecting (or supporting, for that matter) any particular flavor of theory as to the nature of knowledge. Our claim is neither epistemological nor metaphysical, but methodological. And it is our contention that even the contextualist faces serious problems when attempting to use intuitions as evidence. Let us distinguish between *contextualism* simpliciter, which is the view as quoted above; and *armchair contextualism*, which is contextualism conjoined with the further claim that the truth conditions of ‘S knows that p’, including their dimensions of contextual variation, can be determined by philosophers’ appeal to intuitions without any careful scientific study. While we have no quarrel with contextualism itself, we would take armchair contextualism to be as much one of our targets as all other forms of intuition-driven philosophy.

First, it is not clear how easily the armchair contextualist can explain our results. Contextualists have some freedom when it comes to defining the scope of contexts, in this case regarding the relevant time frame for determining contexts. One option is to narrow the time frame for contexts. For instance, Lewis’ principle of accommodation maintains that interlocutors will continually adjust their standards for attributing knowledge to accommodate those standards

presupposed by *the most recent* conversational moves made by their fellow interlocutors. Yet our results indicate that the willingness of subjects to attribute knowledge to Charles is responsive to more than just what immediately precedes the Truetemp Case. Thus, to accommodate our data, contextualists of any stripe will need to widen the time frame for contexts. But this move creates problems for the armchair contextualist, by making it methodologically intractable to determine when any two persons are in the same context. Consider the fact that subjects who are given the same survey version, such that they encounter the thought-experiments in the same order, still have widely varying responses to each case; their attitudes toward a statement attributing knowledge to Charles cover the entire Likert Scale, from strongly agree (5) to strongly disagree (1) in every version of our survey. If armchair contextualists attribute this instability to differences in context, they will have to widen the time frame for contexts beyond the scope of our survey, such that what the subjects encountered before taking our survey also determines the context in which they attribute knowledge in the survey cases. This suggests that it would be impossible to establish that two people are in the same context. In short: if they are not willing to consider contexts to be sufficiently wide, then they cannot begin to accommodate our data; and once they begin to go wide enough, it seems that there will be no armchair-specifiable way to determine just where contexts start and stop. (The temporal extent of contexts is a good candidate for an interesting *empirical* question, though!)²⁹

²⁹ Although we are not here disputing contextualism simpliciter, we would note that it is also an interesting empirical question just how good well contextualism about 'knows' can explain the phenomena. As an anonymous referee for this journal suggests, contextualism about knowledge-attributions might be ruled out by checking whether these effects occur in non-epistemic cases.

Second, if the armchair contextualist explains the instability of intuitions about the Truetemp Case in terms of a relevant shift in context between various versions of our survey—a shift in context based on whether, and which, other thought-experiments have been presented first—a question arises regarding the Fakebarn Case. Responses to the Fakebarn Case are relatively stable with regard to whether, and which, other thought-experiments are presented before it, which seems *prima facie* difficult for the armchair contextualist to accommodate. What the armchair contextualist must do is provide a legitimate (and armchair-accessible) reason for thinking that presenting a clear case of knowledge before presenting the Truetemp Case represents a relevant shift in context that accordingly shifts the standards for knowledge attribution, whereas presenting the same clear case of knowledge before presenting the Fakebarn Case does not represent a relevant shift in context, and therefore does not shift the standards for knowledge attribution. We are pessimistic as to the armchair contextualists ability to discharge this theoretical burden.

4. Conclusion

Our results build on an existing body of empirical research demonstrating that intuitions vary according to factors irrelevant to the issues thought-experiments are designed to address. Specifically, we found that intuitions about the Truetemp Case vary depending on whether, and which, other cases are presented before it. Such variability calls into question the legitimacy of using the intuitions generated by the Truetemp Case as evidence against reliabilism. But it is unclear what about this case makes it susceptible to these effects, which raises questions about the reliance on intuitions about thought-experiments more generally, especially given that this is not the only case called into question by empirical research. We take the growing body of

empirical data impugning various intuitions to present a real challenge for philosophers who wish to rely on intuitions as evidence. We would structure our argument here in conditional terms. First, we suggest that intuition-deploying philosophers need to recognize that if their intuitions are sensitive to variables irrelevant to the issues thought-experiments are designed to address, then they are ill-suited to do the work philosophers ask of them. Second, the status of the antecedent of that conditional is an empirical one, and we take ourselves to have offered here some initial evidence in its favor. We certainly do not take ourselves to have offered anything like a general proof of the unreliability of all intuitions (nor do we think that any such proof would be either possible or desirable). But we do take ourselves to have raised a serious empirical worry that philosophers need to begin deciding how to address. Ceasing all philosophical intuiting at this point may well be intellectually premature; but ignoring the worry that we have raised would likewise be intellectually irresponsible. We propose that philosophers who wish to continue relying on intuitions as evidence begin empirically investigating intuitions about their favorite thought-experiments to determine whether, and which, intuitions may be taken as evidence. Perhaps they will find that, *contra* the worries we have raised here for the likes of Truetemp, there is nothing amiss with their own preferred intuitions. But our results should make live and salient the possibility that they will find that their practice may in fact be built on an unacceptably shifting foundation.

Appendix 1: Epistemological Intuitions Survey

We are investigating what different people's opinions are about knowledge. In each question, please indicate to what extent you agree or disagree with that statement.

1. Dave likes to play a game with flipping a coin. He sometimes gets a "special feeling" that the next flip will come out heads. When he gets this "special feeling", he is right about half the time, and wrong about half the time. Just before the next flip, Dave gets that "special feeling", and the feeling leads him to believe that the coin will land heads. He flips the coin, and it does land heads.

Please indicate to what extent you agree or disagree with the following claim: "Dave knew that the coin was going to land heads."

___ Strongly agree ___ Agree ___ Neutral ___ Disagree ___ Strongly disagree

2. One day Charles was knocked out by a falling rock; as a result his brain was "rewired" so that he is always right whenever he estimates the temperature where he is. Charles is unaware that his brain has been altered in this way. A few weeks later, this brain rewiring leads him to believe that it is 71 degrees in his room. Apart from his estimation, he has no other reasons to think that it is 71 degrees. In fact, it is 71 degrees.

Please indicate to what extent you agree or disagree with the following claim: "Charles knows that it is 71 degrees in his room."

___ Strongly agree ___ Agree ___ Neutral ___ Disagree ___ Strongly disagree

3. Suzy looks out the window of her car and sees a barn near the road, and so she comes to believe that there's a barn near the road. However, Suzy doesn't realize that the countryside she is driving through is currently being used as the set of a film, and that the set designers have constructed many fake barn facades in this area that look as though they are real barns. In fact, Suzy is looking at the only real barn in the area.

Please indicate to what extent you agree or disagree with the following claim: "Suzy knows there is a barn near the road."

___ Strongly agree ___ Agree ___ Neutral ___ Disagree ___ Strongly disagree

4. Karen is a distinguished professor of chemistry. This morning, she read an article in a leading scientific journal that mixing two common floor disinfectants, Cleano Plus and Washaway, will create a poisonous gas that is deadly to humans. In fact, the article is correct: mixing the two products does create a poisonous gas. At noon, Karen sees a janitor mixing Cleano Plus and Washaway and yells to him, "Get away! Mixing those two products creates a poisonous gas!"

Please indicate to what extent you agree or disagree with the following claim: "Karen knows that mixing these two products creates a poisonous gas."

___ Strongly agree ___ Agree ___ Neutral ___ Disagree ___ Strongly disagree