Running Head: ON INTUITIONAL STABILITY

On intuitional stability: The clear, the strong, and the paradigmatic

Jennifer Cole Wright

College of Charleston

Correspondence concerning the article should be addressed to Jennifer Cole Wright, Department of Psychology, College of Charleston, 57 Coming Street, Charleston, SC 29424, email: wrightjj1@cofc.edu.

Word Count: abstract – 103 words; entire document (including abstract) – 5,640 words

Abstract

Skepticism about the epistemic value of intuition in theoretical and philosophical inquiry has recently been bolstered by empirical research suggesting that people's concrete-case intuitions are vulnerable to irrational biases (e.g., the order effect). What is more, skeptics argue that we have no way to "calibrate" our intuitions against these biases and no way of anticipating intuitional instability. This paper challenges the skeptical position, introducing data from two studies that suggest not only that people's concrete-case intuitions are often stable, but also that people have introspective awareness of this stability, providing a promising means by which to assess the epistemic value of our intuitions.

On intuitional stability: The clear, the strong, and the paradigmatic

Intuition—what it is and how, when, and why it works—has recently received renewed attention in philosophy, cognitive science, and psychology. There has been much debate concerning the nature of intuition (Audi, 2004; Bealer, 1999, 2000; Claxton, 1998; Huemer, 2005; Kornblith, 1998; Laughlin, 1997; Osbeck, 1999, 2001; Parsons, 2000; Pust, 2000; Sosa, 1999, 2007; Williamson, 2007; Wisniewski, 1998), as well as what sort of cognitive process intuiting might be or involve (Cummins, 1998; Denes-Raj & Epstein, 1994; Dorfman, Shames, & Kilstrom, 1996; Epstein, Lipson, Holstein, & Huh, 1992; Gendler, 2007; Osbeck, 1999; Shafir, 1998; Sloman, 1996). There has also been debate about what role intuitions might play in logic and mathematics (Bealer, 2000; Bonjour, 1998; Casullo, 2003; Parsons, 1986, 2000; Sosa, 2006; Wright, 2004), epistemology (Alexander & Weinberg, 2007; Bealer, 1992; Brown, 2006; Nagel, 2007; Weinberg, 2007; Willamson, 2004), metaphysics (Bealer, 2002, 2004; Bonjour, 1998; Jackson, 1994, 1998; Pust, 2004; Sosa, 2000, 2006), morality (Audi, 2004; Bartsch & Wright, 2005; Dancy, 1991, 2006; Haidt & Joseph, 2004; Huemer, 2005; Jackson, 1998; Macnamara, 1991), and a variety of other areas.[1]

One such debate concerns the *epistemic*[2] status of intuitions. This debate centers around the following question: Is it legitimate, epistemically speaking, for individuals to form beliefs about matters of logic, mathematics, metaphysics, epistemology, morality, etc. on the basis of their intuitions about theoretical principles and/or concrete cases (involving actual or hypothetical examples)? In other words, do intuitions have some positive epistemic value?

While there are many who endorse an affirmative answer to this question (e.g., Bealer, 1992, 1999, 1998, 2000, 2004; Bonjour, 1998; Jackson, 1998; Pust, 2000; D. Sosa, 2006; E. Sosa, 1999, 2006, 2008; Williamson, 2004; cf. Osbeck, 1999, 2001), an increasing number of

philosophers, cognitive scientists, and psychologists express a deep skepticism about intuition's epistemic value (see, e.g., Cummins, 1998; Denes-Raj & Epstein, 1994; Gendler, 2007; Hintikka, 2001; Machery, Mallon, Nichols, & Stich, 2004; Nichols & Knobe, 2007; Nichols, Stich, & Weinberg, 2003; Nisbett, Peng, Choi, & Norenzayan, 2001; Redelmeir & Shafir, 1995; Weinberg, 2007; Weinberg, Nichols, & Stich, 2001). In fact, whereas extreme skepticism about perception and memory might be considered somewhat 'academic' (D. Sosa, 2006), skepticism about intuition is thought by many (e.g., Machery, Mallon, Nichols, & Stich, 2004; Nichols & Knobe, 2007; Nichols, Stich, & Weinberg, 2003) to have serious implications for philosophical methodology.

This skepticism has recently been fortified by empirical research showing that concrete-case intuitions are vulnerable to irrational biases. Swain, Alexander, & Weinberg (2008), for example, found that people's responses to concrete cases were vulnerable to an 'order effect' (Tversky & Kahneman, 1974). Specifically, Swain, et al. (2008) found that participants' concrete-case judgments about the *True-Temp* case (a much discussed thought-experiment in contemporary epistemology in which a man is unwittingly led, through a 'brain rewiring', to form true beliefs about the current temperature; see Lehrer, 1990) were significantly influenced by what case they had previously considered. Participants were more likely to say that True-Temp knew (as opposed to 'merely believed') that the temperature was 71 degrees if they had just previously read the case about Dave, a man who formed the true belief that the next coin he flipped would land heads because he had a 'special feeling' right before he flipped the coin, and they were *less* likely to say that True-Temp knew if they'd just previously read the case about Karen, a woman who formed a true belief about how to create a poisonous gas on the basis of reading an article about it in a top scientific journal.

Based on these findings, Swain, et al. (2008) concluded that, to the extent that people's concrete-case intuitions are influenced by irrational biases such as one's previously elicited intuitions, they do not possess the sort of epistemic status that they have heretofore been taken to possess. They further concluded that the instability of intuitions demonstrated by their study (and others: e.g., Machery, Mallon, Nichols, & Stich, 2004; Weinberg, Nichols, & Stich, 2001) brings into question our reliance on intuitions as sources of evidence for theoretical/philosophical positions, writing 'we contend that this instability undermines the supposed evidential status of these intuitions, such that philosophers [and others] who deal in intuitions can no longer rest comfortably in their armchairs' (2008, 1).

Is such a strong conclusion warranted? Some have argued that it is not, either due to a variety of methodological and conceptual difficulties (none of which will be touched upon here – for a discussion of some of these issues, see Laio, 2008 and elsewhere) or on the grounds that Swain, et al. (2008) hardly provide a definitive demonstration of intuitional instability, having found it in only one particular case. However, the true weight of the Swain, et al. (2008) challenge is not that all (or even most) intuitions are unstable, but rather that we have no way of "calibrating" our intuitions, no way of anticipating the conditions under which our concrete-case intuitions will be vulnerable to irrational biases, such as the order effect (for more on this worry, see Weinberg, 2007). This being the case, an adequate response to Swain, et al.'s challenge needs to do more than simply demonstrate the stability of some (or even most) intuitions – it needs to identify a reliable method by which to track that stability and provide insight into why *certain* intuitions, but not others, are stable. In the absence of this, the epistemic legitimacy of consulting our intuitions remains open to skepticism.

The goal of the two studies reported here was to take up the challenge. Their guiding hypothesis was twofold: 1) that only *some* intuitions (that is, intuitions about *certain sorts* of cases) are vulnerable to intuitional instability and that people are implicitly aware of which cases these will be, and 2) that several potentially reliable methods for tracking intuitional instability exist – among them, the introspectively accessed confidence and belief strength of those doing the intuiting.

Study 1

Methods

Participants

188 undergraduate college students (87 males, 101 females; dominantly Caucasian) from the University of Wyoming participated in this study. Participants were recruited through the Introduction to Psychology research pool and received research credit for their participation. Being dominantly college freshman, the assumption was that the participants had received little to no explicit philosophical training (though this question was not asked).

Materials & Procedure

Participants received a randomized series of the Swain, et al. (2008) cases as 'filler tasks' while participating in one of two larger, unrelated studies. The set included the *True-Temp* case, the *Coin-Flip* case, the *Fake-Barn* case, and the *Testimony* case (see Appendix), all four of which were presented to participants in a counterbalanced order. After reading each case, participants were asked whether the subject in the case *knew* a specific proposition (e.g., for *True-Temp*, whether the temperature was 71 degrees), to which participants answered YES or NO. They were then asked to rate how confident they were about their answer (0 = not very confident to 5 = very confident).

Results

Preliminary note: There were no gender differences found and so all analyses to follow were collapsed across gender.

Swain, et al. (2008) had found that participants were *more* likely to judge that *True-Temp* really knew the temperature if the case was immediately preceded by *Coin-Flip* and *less* likely to do so if it was immediately preceded by *Testimony*. A similar pattern emerged in this study. Examining those cases in which *True-Temp* was the second case participants considered, being directly preceded by one of the other three cases (KTxx, DTxx, or STxx), the results revealed that participants were significantly more likely to attribute knowledge to *True-Temp* when it immediately followed *Coin-Flip* (55%) than when it immediately followed either *Testimony* (40%) or *Fake-Barn* (26%), χ^2 (2, N = 143) = 8.25, p = .016 (Figure 1).

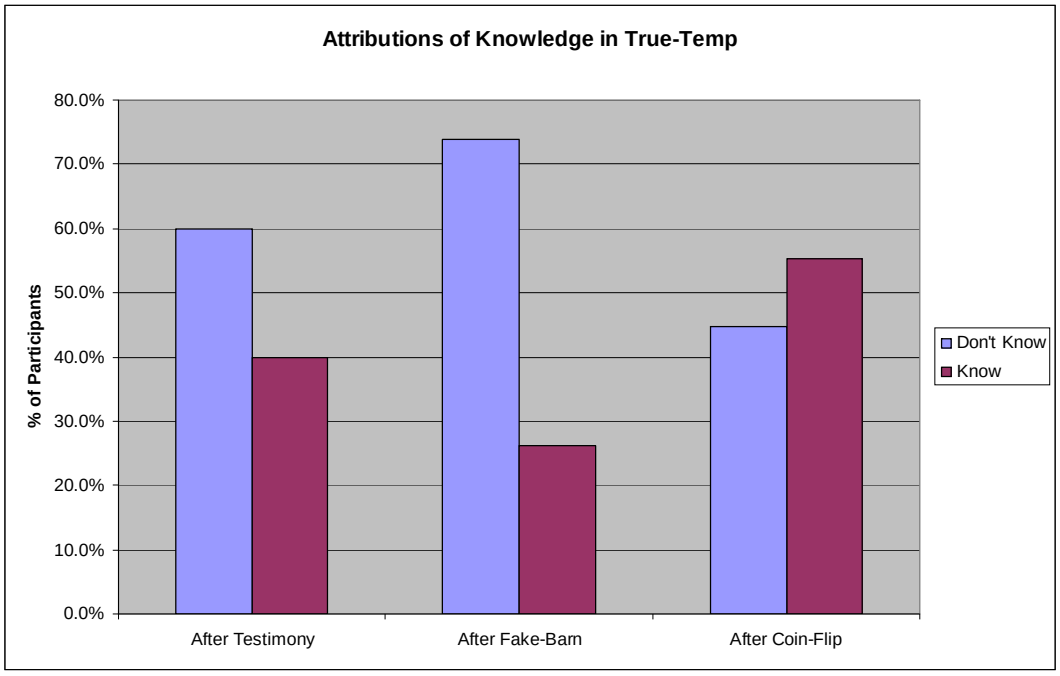

Figure 1. Attributions of Knowledge in True-Temp

A similar trend emerged for *Fake-Barn*: participants were (marginally) less likely to count Suzy's mental state as knowledge when the case immediately followed either *Testimony* or

Coin-Flip (40% and 39%, respectively) than when it immediately followed *True-Temp* (59%): χ^2 (2, N = 144) = 4.91, p = .086 (Figure 2).

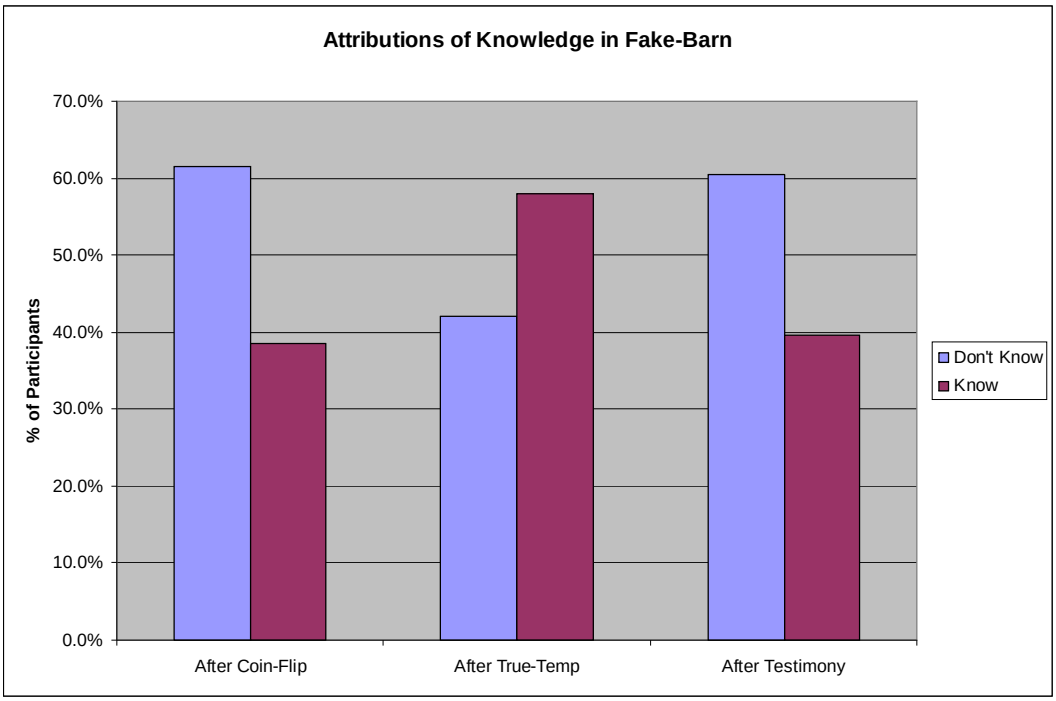

Figure 2. Attributions of Knowledge in Fake-Barn

Participants' judgments about *Testimony* and *Coin-Flip*, on the other hand, did not display vulnerability to the order effect. Participants were equally likely to attribute knowledge in *Testimony*, regardless of which case immediately preceded (79-84%), χ^2 (2, N = 139) = .50, p = .77. Likewise, participants were equally likely to *fail* to attribute knowledge in *Coin-Flip*, regardless of which case immediately preceded (0-6%), χ^2 (2, N = 133) = 2.50, p = .29 (Figure 3).

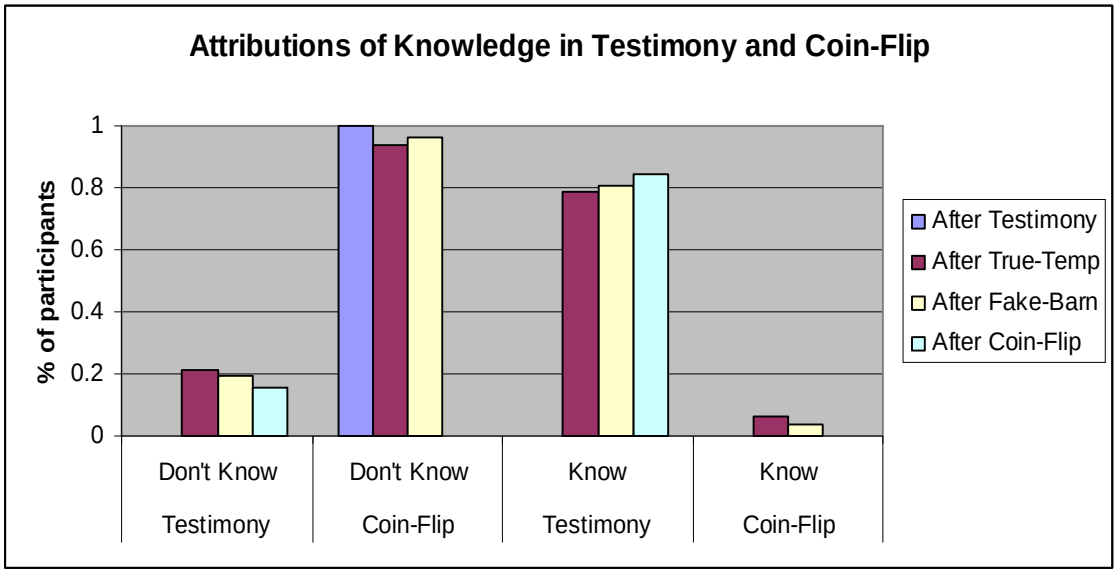

Figure 3. Attributions of Knowledge in Testimony and Coin-Flip

Of central importance is the fact that participants' themselves introspectively tracked this intuitional stability. Paired sample t-tests revealed that participants were significantly more confident in their judgments about *Coin-Flip* (M = 4.4, SE = .06) and *Testimony* (M = 4.5, SE = .06) than they were in their judgments about *True-Temp* (M = 3.9, SE = .09) and *Fake-Barn* (M = 3.9, SE = .08), ts(187) = 5.4 to 6.7, ps < .001, while their confidence did not significantly differ between the two stable and two unstable cases[3], t(187) = 1.6, p = .11 and t(187) = .27, p = .79, respectively.

And this was true regardless of order in which the cases were presented. Participants expressed higher levels confidence in their judgments for *Coin-Flip* and *Testimony* than for *True-Temp* and *Fake-Barn* whether they were the first cases considered (M = 4.6, SE = .17 and M = 4.4, SE = .12 vs. M = 3.7, SE = .17 and M = 3.8, SE = .17, respectively) or the last cases considered (M = 4.5, SE = .12 and M = 4.5, SE = .12 vs. M = 3.9, SE = .20 and M = 3.9, SE = .18, respectively, Figure 4). Thus, it would appear that participants' confidence served as an introspective indicator of intuitional stability.

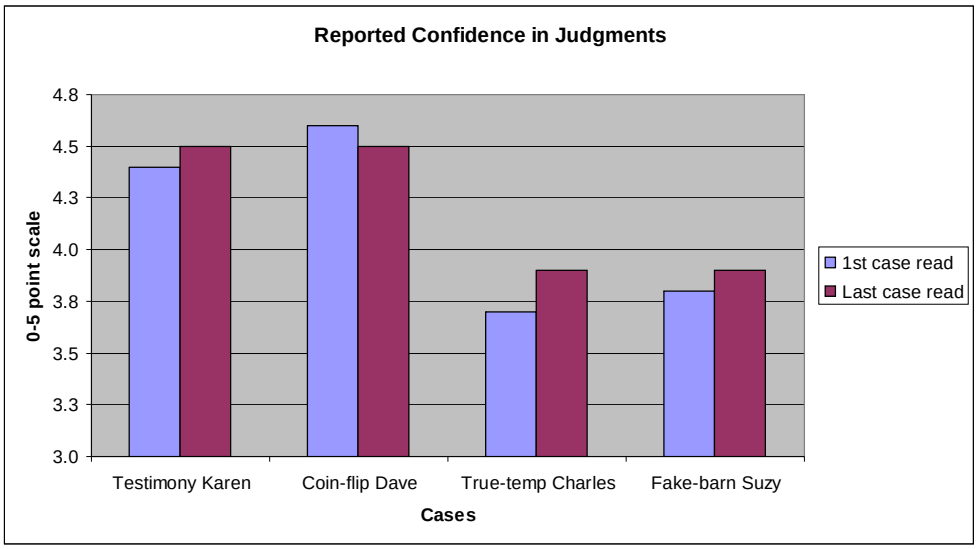

Figure 4. Reported Confidence in Judgments

To further explore the relationship between stability and confidence, nominal logistic regressions were performed to see if participants' confidence levels could be used to predict which case was being considered. Participants' confidence was regressed as a covariate over the cases (1= *True-Temp*, 2 = *Fake-Barn*, 3 = *Coin-Flip*, 4 = *Testimony*), with each case functioning as the reference case. Each model revealed confidence to be a strong predictor of stable vs. unstable cases, $X^2(3, N=188) = 26.8$, $p < .001$.

Specifically, the results show that confidence was a significant predictor of whether the case being considered was stable (*Testimony* or *Coin-Flip*) or unstable (*True-Temp* or *Fake-Barn*). For every 1 unit increase in participants' confidence, the odds of the case being *Testimony* (over *True-Temp*) increased by 203% (or a factor of 2.03), $X^2(1) = 9.6$, $p = .002$, and the odds of the case being *Coin-Flip* (over *True-Temp*) increased by 273%, $X^2(1) = 15.4$, $p < .001$. Changes in confidence did not distinguish between *True-Temp* and *Fake-Barn*, $X^2(1) = .134$, $p = .714$. Likewise, for every 1 unit increase in participants' confidence, the odds of the case being *Fake-*

Barn (over *Testimony*) decreased by 53%, $X^2(1) = 7.7$, $p = .005$, but changes in confidence did not distinguish between *Testimony* and *Coin-Flip*, $X^2(1) = 1.24$, $p = .266$.

Framed in terms of probability, as participants' confidence increased, the probability that the case being considered was either *Testimony* or *Coin-Flip* increased significantly, from around 5% at a confidence of '1' to about 30% (*Testimony*) to 60% (*Coin-Flip*) at a confidence of '7', and the probability that the case being considered was *True-Temp* or *Fake-Barn* decreased significantly, from about 40% (*Fake-Barn*) to 50% (*True-Temp*) at '1' to around 5% at '7'. The case probabilities merged and became approximately equal at a confidence level of '4' (Figure 5: INSERT PLOT1 HERE).

Discussion

In line with Swain, et al. (2008), this first study found participants' knowledge attributions in the *True-Temp* case to be unstable and their fluctuation to be dependent upon which case had been previously considered. The same was true for the *Fake-Barn* case. And, also in line with Swain, et al., both *Testimony* and *Coin-Flip* remained stable across the counterbalanced order in which they were presented.

Importantly, participants seemed to be introspectively tracking this instability, reporting significantly more confidence when considering cases that elicited stable judgments than when considering cases that elicited unstable judgments, regardless of order of presentation. As the results from the logistic regressions reflect, the lower participants' confidence, the more likely the case being considered displayed instability – the higher their confidence, the more likely it displayed stability.

Why would participants be vulnerable to bias in only two of the four cases? One reasonable explanation is that, when considering *True-Temp* and *Fake-Barn,* people's intuitions

about them were less clear. If so, then it would make sense for participants to turn elsewhere, such as to the case that they had just previously considered, for information that would help to determine their judgment. For example, when considering whether or not *True-Temp*'s mental state should count as knowledge, perhaps participants who saw *Coin-Flip* first were more inclined to say 'yes' because it looks a lot *more* like knowledge than a 'special feeling'. On the other hand, it looks a lot *less* like knowledge than testimony from a top scientific journal. In short, under circumstances where our intuitions are not as clear, it would be natural for us to bring other information to bear on our judgments.

This could also explain why participants' knowledge attributions were not influenced by the preceding cases for either *Testimony* or *Coin-Flip*. As Swain, et al. (2008) noted, these cases are 'clear cases' – or what we might call *paradigmatic* cases – of having (or failing to have) knowledge. Arguably, *Testimony* is the sort of case that most people (people with a reasonable degree of conceptual competence[4]) would agree is an instance of 'knowledge', just as *Coin-Flip* is the sort of case that most would agree *fails* to be an instance of 'knowledge'. It is not surprising to find that people's intuitions stabilize around paradigmatic cases – cases that are clear instances of our concepts – and so are not vulnerable to the sorts of biasing factors that Swain and colleagues argue undermines intuition's evidential status.

In order to investigate intuitional stability in more depth, a second study was conducted that expanded upon Study 1 in three ways. First, it introduced additional cases for participants to consider; second, it introduced belief strength as an additional introspective measure of stability. Research on attitude and belief strength has found strongly held beliefs to be more stable over time, more resistant to change, and less sensitive to contextual influences (for reviews see Petty

& Krosnick, 1995; Krosnick & Petty, 1995). Thus it was hypothesized that belief strength might serve as another good introspective indicator of stability.

Lastly, perceived consensus was introduced as a (rough) proxy for 'paradigmaticity'. The hypothesis here was that participants would be likely to view clear (paradigmatic) cases as the sorts of cases everyone would agree upon – therefore, the more paradigmatic the case under consideration, the higher the degree of peer consensus they should report. The claim is not that perceived consensus would serve as a measure of *actual* paradigmaticity (which may involve factors outside of mere agreement), but rather that would provide important insight into how paradigmatic people *perceive* the cases they are considering to be.

Study 2

Methods

Participants

181 undergraduate college students (33 males, 148 females; dominantly Caucasian) from the College of Charleston participated in this study. Participants were recruited through the Introduction to Psychological Science research pool and received research credit for their participation. 93% of the participants had no philosophical training, 6% had taken or were currently were enrolled in Introduction to Philosophy, and 1% in some other undergraduate philosophy course.

Materials & Procedure

This time participants were presented with three different sets of cases, nine cases in total (see Appendix). Two of the sets involved cases in epistemology, expanding upon the cases considered in Study 1 (Set 1: *Perception*, *True-Temp*, and *Coin-Flip*; Set 2: *Testimony*, *Farmer*, and *Guess*) and one set involved cases in ethics[5] (Set 3: *Break-Promise*, *Hide-Bombers*, *Sell-*

iPod). Once again, the cases were presented to the participants in a counterbalanced order, though this time the counterbalancing occurred both *within* sets (e.g., *Perception*/*True-Temp*/*Coin-Flip*, *True-Temp*/*Coin-Flip*/*Perception*…) and *between* sets (e.g., Set1/Set2/Set3, Set2/Set3/Set1…). It was anticipated that six of these cases (*Perception*, *Coin-Flip*, *Testimony*, *Guess*, *Break-Promise*, and *Sell-iPod*) would elicit stable judgments and the other three would elicit unstable judgments.

After reading an epistemology case, participants were asked whether the subject in the case *knew* a specific proposition, to which participants answered YES or NO. After reading an ethics case, they were asked whether the action performed in the case was *morally wrong*, to which they answered YES or NO. Once again, participants were asked to rate on a Likert scale (this time, a 7-point scale in order to provide a neutral midpoint) how confident they were about their answer (1 = not very confident to 7 = very confident). Participants were also asked to rate on a 7-point Likert scale how strongly they believed their answer (1 = not very strongly to 7 = very strongly). Finally, participants were asked a perceived consensus question: If 100 other College of Charleston students were asked the same question, how many do you think would give the same answer you did? (1 = none of them to 7 = all 100 of them). The order of these questions was counterbalanced between participants.

Results

Preliminary note: There were no gender differences or differences between participants with vs. without philosophical training, so analyses reported below were collapsed across these groups. In addition, all analyses conducted with participants' confidence in Study 1 were replicated and participants' confidence and belief strength ratings were highly correlated across

all nine cases (*r*s =.86 to .99, *p*s < .001), so for the sake of brevity analyses with confidence are not reported below.

Of the nine cases that participants considered, six (as anticipated) elicited stable intuitions. For *Perception*, *Coin-Flip*, *Testimony*, *Guess*, *Sell-iPod*, and *Break-Promise*, the order of presentation did not matter. Participants dominantly attributed knowledge in *Perception* (80-90%) and *Testimony* (84-87%) and failed to attribute knowledge in *Coin-Flip* (3%) and *Guess* (0-7%), regardless of order. Participants also dominantly judged the action to be morally wrong in *Sell-iPod* (100%) and not wrong in *Break-Promise* (100-97%), regardless of order. Examining the pattern of participants' answers when each of these cases was immediately preceded by the other cases in its set revealed no significant variation for any of them, X^2s(1, *N*s = 57-61) = 0.0 to 1.05, *p*s = .31-.99. All of these cases were perceived by participants as being highly paradigmatic (in the sense that participants reported a high degree of agreement in their peers): *M*s = 5.6 – 6.5 (SEs = .06-.09).

Two of the remaining cases elicited unstable judgments: *True-Temp* and *Hide-Bombers*. Examining the cases in which *True-Temp* was directly preceded by one of the other two cases, the results showed that participants were significantly more likely to say that *True-Temp* knew the temperature immediately after reading *Coin-Flip* (84%) than after reading *Perception* (57%), $X^2(1, N = 61) = 5.4$, $p = .020$. Likewise, when reading *Hide-Bombers*, participants were marginally more likely to say that what Hilda did was morally wrong immediately after reading *Sell-iPod* (55%) than after reading *Break-Promise* (32%), $X^2(1, N = 60) = 3.2$, $p = .073$ (Figure 6). These two cases were seen as significantly less paradigmatic than either the stable 'yes' or the stable 'no' cases: *M*s 4.9 and 5.2 (*SE*s = .08-.09), *t*s(174-179) = 6.4-13.3, *p*s < .001.

The final case, *Farmer*, was an interesting case. Examining participants' knowledge attributions revealed that it was not unstable, in the sense of demonstrating an order effect, $X^2(1, N = 61) = .008$, $p = .93$, but neither was it paradigmatic – as with the unstable cases, participants were strongly divided over whether or not the Farmer knew his cow was in the field, approximately 1/3rd saying he did know and 2/3rd saying he did not (Figure 6).

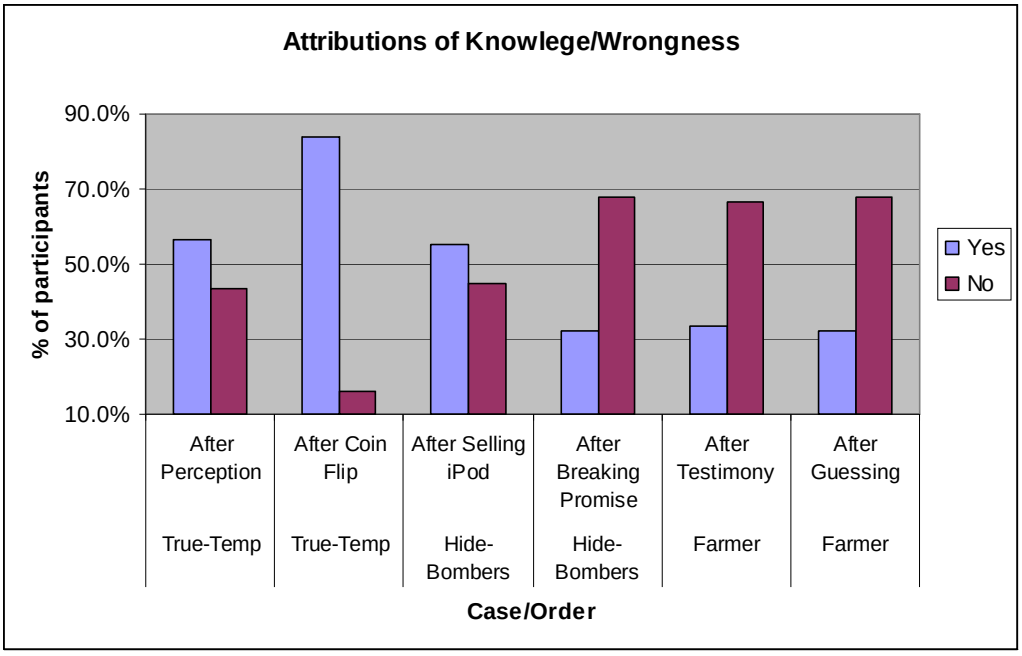

Figure 6. Attributions of Knowledge/Wrongness in *True-Temp*, *Hide-Bombers*, and *Farmer*

This fact was nicely reflected in participants' reports of paradigmaticity, which for *Farmer* fell significantly *in between* their reported paradigmaticity for the stable vs. unstable epistemology cases: paired-sample t-tests showed perceived consensus for *Farmer* ($M = 5.3$, $SE = .09$) to be lower than their perceived consensus for the stable-yes (*Perception*/*Testimony*) and stable-no (*Coin-Flip*/*Guess*) epistemology cases (Ms = 6.0/5.9 and 5.6/6.4, SEs =.07-.09, respectively), ts(178) = 2.0-7.7, ps from .05 to <.001, and yet higher than the unstable (*True-Temp*) epistemology case ($M = 4.9$, $SE = .08$), $t(177) = 5.2$, $p < .001$ (Figure 7).

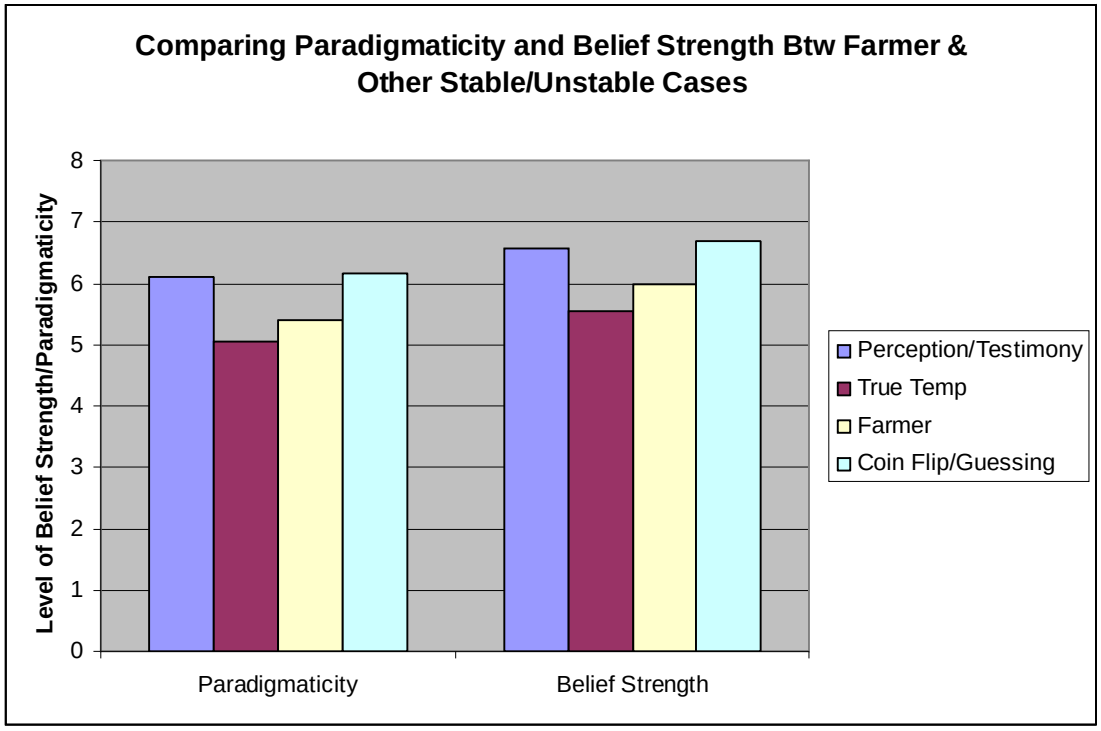

Figure 7. Paradigmaticity & Belief Strength between Farmer vs. Stable/Unstable Cases

Interestingly, participants' belief strength showed a similar pattern: paired-sample t-tests revealed that participants' belief strength for *Farmer* (M = 5.9, SE = .09) was significantly lower than for the stable-yes (*Perception*/*Testimony*) and stable-no (*Coin-Flip*/*Guess*) epistemology cases (Ms = 6.5/6.3 and 6.4/6.7, SEs =.05-.07, respectively), ts(178) = 3.5-9.2, ps =<.001, but also significantly higher than for the unstable (*True-Temp*) case (M = 5.4, SE = .10), t(178) = 5.3, p <.001 (Figure 7). Given that participants' introspective judgments appear to be locating *Farmer* in between the stable and unstable cases, it will be heretofore referred to as an "intermediate" case.

More generally, a within-subjects ANOVA with stability (stable-yes, unstable/ intermediate, stable-no) and set (epist1, epist2, ethics) as within-subjects factors revealed that participants' belief strength was significantly higher for the stable cases than for the unstable/intermediate cases across all three sets, $F(2,346) = 87.7$, $p < .001$, $\eta^2 = .34$.

Importantly, this suggests that the same general relationship between stability and confidence also holds for stability and belief strength – and that it does so across multiple sets of

cases, both epistemological and ethical. But does this mean that belief strength, like confidence, can be used to predict case? To investigate this, nominal logistic regressions (separate for each set) were performed with belief strength as the covariate over case (Set 1: 1 = *Perception*, 2 = *Coin-Flip*, 3 = *True-Temp*; Set 2: 1 = *Testimony*, 2 = *Guess,* 3 = *Farmer*; Set 3: 1 = *Sell iPod*, 2 = *Break-Promise,* 3 =*Hide-Bombers*).

For Set 1, belief strength was a strong predictor of stable vs. unstable cases, $X^2(2, N=179) = 62.1$, $p < .001$. For every 1 unit increase in participants' belief strength, the odds of the case being *Perception* (over *True-Temp*) increased by 302% (or a factor of 3.02), $X^2(1) = 31.7$, $p < .001$, and the odds of the case being *Coin-Flip* (over *True-Temp*) increased by 272%, $X^2(1) = 28.7$, $p < .001$. As expected, belief strength was not predictive between *Perception* and *Coin-Flip*, $X^2(1) = .30$, $p = .584$.

Framed in terms of probability, as participants' belief strength increased, the probability that they were considering *True-Temp* dropped significantly, from almost 100% at a belief strength of '1' to around 10% at a belief strength of '7'. Likewise, as belief strength increased, the probability that they were considering either *Perception* or *Coin-Flip* increased, from almost 0% at '1' to around 45% at '7'. The case probabilities merged and became approximately equal between '5' and '6' (Figure 8: INSERT PLOT2 HERE).

For Set 2, belief strength was a strong predictor of all three cases, $X^2(2, N=180) = 31.5$, $p < .001$. For every 1 unit increase in participants' belief strength, the odds of the case being *Testimony* (over *Farmer*) increased by 148%, $X^2(1) = 6.7$, $p = .010$, and the odds of the case being *Guess* (over *Farmer*) increased by 292%, $X^2(1) = 20.7$, $p < .001$. And for every 1 unit increase in participants' belief strength, the odds of the case being *Guess* (over *Testimony*) increased by 197%, $X^2(1) = 8.2$, $p = .004$.

As participants' belief strength increased, the probability that they were considering *Farmer* dropped significantly, from around 85% at a belief strength of '1' to around 20% at a belief strength of '7'. Likewise, as belief strength increased, the probability that they were considering *Guess* or *Testimony* increased, from around 0% (*Guess*) to 15% (*Testimony*) at '1' to around 30% (*Testimony*) to 50% (*Guess*) at '7'. The case probabilities in this set never fully merged (Figure 9: INSERT PLOT3 HERE).

Finally, for Set 3, belief strength was once again a strong predictor of stable vs. unstable cases, $X^2(2, N=176) = 73.8$, $p < .001$. For every 1 unit increase in participants' belief strength, the odds of the case being *Sell-iPod* (over *Hide-Bombers*) increased by 903%, $X^2(1) = 21.0$, $p < .001$, and the odds of the case being *Break-Promise* (over *Hide-Bombers*) increased by 411%, $X^2(1) = 21.3$, $p < .001$. Belief strength was not predictive between *Sell-iPod* and *Break Promise*, $X^2(1) = 2.5$, $p = .110$. As participants' belief strength increased, the probability that they were considering *Hide-Bombers* decreased, from almost 100% at a belief strength of '1' to around 10% at '7', while the probability of the case being either *Sell-iPod* or *Break-Promise* increased from almost 0% at '1' to around 45% at '7'. The case probabilities merged and became approximately equal around '6' (Figure 10: INSERT PLOT4 HERE).

Participants' perception of paradigmaticity was also predictive of case in all three sets. In Set 1, every 1 unit increase in peer consensus increased the odds that the case being *Perception* (over *True-Temp*) by 470%, $X^2(1) = 37.3$, $p < .001$, and increased the odds that the case was *Coin-Flip* (over *True-Temp*) by 195%, $X^2(1) = 10.5$, $p = .001$. It also decreased the odds that the case being considered was *Coin-Flip* (over *Perception*) by 42%, $X^2(1) = 15.7$, $p < .001$.

In Set 2, every 1 unit increase in consensus increased the odds of the case being *Testimony* (over *Farmer*) by 162%, $X^2(1) = 10.0$, $p = .002$, and increased the odds of the case

being *Guess* (over *Farmer*) by 238%, $X^2(1) = 22.6$, $p < .001$. It also increased the odds that the case being considered was *Testimony* (over *Guess*) by 147%, $X^2(1) = 4.7$, $p = .035$.

Finally, in Set 3 every 1 unit increase in consensus increased the odds of the case being *Sell-iPod* (over *Hide-Bombers*) by 353%, $X^2(1) = 29.2$, $p < .001$, and increased the odds of the case being *Break-Promise* (over *Hide-Bombers*) by 413%, $X^2(1) = 32.6$, $p < .001$. Consensus was not predictive between *Sell-iPod* and *Break-Promise*, $X^2(1) = .44$, $p = .51$.

On a final note, when either confidence or belief strength are entered into logistic regression equations alongside paradigmaticity (multicollinearity issues make it problematic to enter confidence and belief strength together), each variable remains predictive of stable vs. unstable cases in Sets 1 and 3 (*p*s between < .001 and .042), while only paradigmaticity is significant in Set 2 ($p = .001$ for paradigmaticity, $p = .20$ for confidence; $p = .019$ for paradigmaticity, $p = .11$ for belief strength), perhaps because Set 2 contains the intermediate case. Thus, paradigmaticity, on the one hand, and confidence/belief strength, on the other, appear to be independently predictive of intuitional instability.

General Discussion

The worry introduced by Swain, et al. (2008) is that philosophers' reliance on intuitions in argumentation for/against particular theses and theoretical positions is problematic because at least some of those intuitions are epistemically vulnerable to irrational biases like the order effect. This worry gets its teeth not primarily because of the sheer number of intuitions that could be vulnerable, but more importantly because of our supposed inability to anticipate (and protect against) this vulnerability. The thought is that we lack reliable methods by which to track intuitional instability – and, therefore, we cannot know when our intuitions are being negatively impacted by it and when they are not.

Contrary to this claim, the studies reported here suggest that there *are* ways for us to anticipate intuitional instability – in fact, several different (though related) ways. For one, the participants in both studies clearly experienced more confidence in their answers when considering stable cases than when considering unstable cases, regardless of whether those cases involved the application of epistemological or ethical concepts. In addition, participants had significantly stronger beliefs (or, held their beliefs more strongly) about the stable cases than about the unstable cases.

Indeed, both confidence and belief strength were good predictors of whether the case being considered was stable or unstable. This is interesting because while confidence seems a more purely cognitive construct – the degree to which you experience conceptual clarity or certainty – belief strength is often treated as more of an affective construct, sometimes being employed in research paradigms as a proxy for things like 'attitude extremity' and 'emotional intensity' (e.g., Wright, Cullum, & Schwab, 2008). And it seems possible for a person to have a high degree of confidence in a belief they don't hold particularly strongly: my son and I could both be highly confident in our shared belief that the Pittsburg Steelers emerged as the best team in the NFL in 2009, but because he is such an avid football fan, he might hold that belief much more strongly (in the sense that it would be more important to him, etc.) than I. It also seems that people could be *very confident in x* and *believe very strongly in x* for different reasons: someone could be very confident in his belief that our country should make healthcare reform its top priority because of a variety of expert financial analyses he'd read, but then hold this belief very strongly because of a political/philosophical belief that people deserve equal access to healthcare. These potential differences between the two constructs aside, both clearly (and similarly) tracked with intuitional stability.

As mentioned above, one important factor behind the stability of particular cases may be their relative paradigmaticity – that is, the degree to which they represent clear instances of the concept(s) in question. In the epistemological cases participants considered, the concept in question was 'knowledge': whether the subjects *knew* or *didn't know* some particular proposition. And some things strike us as clear examples of knowledge (e.g., beliefs gained through direct perceptual observation under ideal conditions), while some things strike us as clearly *not* knowledge (e.g., randomly guessing the answer to a question, even if you happened to guess the correct answer). In the ethical cases participants were asked whether or not an action was *morally wrong*. Here again, some things strike us as clearly morally wrong (e.g., stealing something of value that is not yours for frivolous reasons), just as some things strike us as clearly *not* morally wrong (e.g., breaking a promise involving a minor obligation because a loved one is in danger). When considering cases such as these, it is less likely that outside factors, such as cases previously considered, will influence our judgments than when we consider difficult borderline (or just otherwise confusing) cases.

Importantly, participants displayed awareness of paradigmaticity (at least insofar as perceived consensus serves as a legitimate measure). The degree of peer consensus that participants reported was strongly related to case stability and participants' peer consensus reports accurately identified between the stable, intermediate, and unstable cases. The stable cases were viewed as being the sorts of cases everyone would agree upon, whereas the unstable and intermediate cases were viewed as more open for disagreement (the unstable cases even more so than the intermediate case).

Collectively, these results suggest two things. First, they suggest that people are able to introspectively track – and thus potentially protect against – their vulnerability to (at least some

forms of) bias. If those intuitive judgments people feel less confident and strongly about are more vulnerable to potentially biasing "outside influences" (or if those cases that are vulnerable to bias are cases that people feel less confident and strongly about), then we can take care with the circumstances under which we elicit intuitions about those cases; we can control exposure to potentially biasing influences. We can also treat such intuitive judgments with caution, granting them less epistemic weight in philosophical/theoretical discourse. Importantly, I would argue that most philosophers and scientists already do this, treating clear/strong intuitions (especially their own) more seriously than unclear/weak ones. Perhaps such efforts could be made more thoughtful and explicit, but I doubt that this would require any drastic changes to current philosophical/theoretical practice.

Second, the results suggest that the more paradigmatic the case, the less vulnerable it will be to (at least some forms of) bias. Again, this seems relatively unsurprising. Clear cases of any particular concept are precisely that: *clear* cases. Our judgments about them, barring substantial conceptual revolution, are unlikely to change. Of course, philosophy is often most interesting (and of most value) when it is working "at the margins", wrestling with unclear and borderline cases. And this raises Swain, et al.'s (2008) worry once again: does this render a potentially extensive area of philosophical debate epistemically vulnerable, philosophers being unable to rely on their intuitions without worry of bias?

This is an important consideration (especially when difficult, non-paradigmatic cases are often where philosophers' rely most heavily on their intuitions), but it seems unlikely to represent an insurmountable problem for several reasons. First, philosophers clearly can (and do) have clear/strong intuitions about non-paradigmatic cases. Such cases, while perhaps vulnerable to bias for the general population of reasonably conceptually competent people (such as the

participants in the studies reported here), may nonetheless be stable for most philosophers. After all, philosophers receive extensive training designed specifically to refine and enhance their conceptual mastery. Such training gives philosophers a greater capacity for discrimination when it comes to concept application (e.g., whether something counts as an instance of knowledge) and, therefore, they may be able to see difficult cases more clearly, and more difficult cases clearly, than the philosophical novice – not unlike learning the difference between *Quercus rubra* (Northern Red Oak) and *Quercus alba* (Pin Oak), which thereafter gives one the ability to distinguish between trees that before that seemed indistinguishable.

What is more, the very process of engaging in philosophical discourse may ultimately *generate* new conceptual clarity where before there was conceptual confusion.[7] That is, the practice of philosophical and theoretical discourse itself may expand and refine our range of conceptual competence, both because of the learning that occurs in the individual and also because of the collective advancement that results for the discipline.

Of course, it is important to note that what confidence and belief strength track with is intuitional stability – not intuitional *accuracy*. Research suggests that people are notoriously overconfident in their judgments across a wide variety of contexts (Arkes, 2001; Einhorn & Hogarth, 1978; Kahneman & Klein, 2009). And, more to the point, simply having clear/strong intuitions does not guarantee that they are also veridical: intuitions are not infallible (and few, if any, philosophers think that they are). Thus, we must be careful not to mistake high degrees of confidence/belief strength as being indicators that we've gotten it right.

What is more, every scientist and philosopher must at some point encounter the line between looking for the theory that best fits one's data and looking at the data in a way that best fits one's theory. The latter is always a danger – and, likewise, there is always the danger that

people's intuitions will be biased by the very training and theoretical advancement that resulted in their heightened conceptual clarity. But, this was not the problem for philosophical intuition that was raised by Swain, et al. (2008) – and thus, not the problem the studies reported here were designed to address. And the good news is that whatever epistemically suspect reasons (e.g., unwarranted theoretical commitments) for intuitional stability that may exist, intuitional instability is *one* worry that we don't need to be too worried about.

References

Alexander, J., & Weinberg, J. (2007). Analytic epistemology and experimental philosophy, *Philosophy Compass*, 2:1, 56-80.

Arkes, H. R. (2001). Overconfidence in judgmental forecasting. In J. S. Armstrong (Ed.), *Principles of forecasting: A handbook for researchers and practitioners* (pp. 495-516). Boston: Kluwer Academic.

Audi, R. (2004). *The Good and the Right: A Theory of Intuition and Intrinsic Value*. Princeton University Press.

Bartsch, K., & Wright, J.C. (2005). Towards an intuitionist account of moral development [commentary]. *Behavioral and Brain Sciences*, 28, 546-547.

Bealer, G. (1992). The incoherence of empiricism. *Aristotelian Society: Supplementary Volume*, 66, 99-138.

Bealer, G. (1999). Intuition and the autonomy of philosophy. In M. DePaul and W. Ramsey (eds.), *Rethinking Intuition: The Psychology Of Intuition And Its Role In Philosophical Inquiry* (pp. 201-239). Lanham: Rowman and Littlefield.

Bealer, G. (2000). A theory of the a priori. *Pacific Philosophical Quarterly,* 811, 1-30.

Bealer, G. (2002). Model epistemology and the rationalist renaissance. In T. Gendler (ed), *Conceivability and Possibility* (pp 71-125). Oxford: Oxford University Press.

Bealer, G. (2004). The origins of modal error. *Dialectica: International Journal of Philosophy of Knowledge,* 581, 11-42.

Bonjour, L. (1998). *In Defense of Pure Reason*. Cambridge: Cambridge University Press.

Bowers, K., Farvolden, P., & Mermigis, L. (1995). Intuitive antecedents of insight. In S. Smith and T. Ward (eds.). *The Creative Cognition Approach*. Cambridge: Bradford/MIT Press.

Bowers, K., Regehr, G., & Balthazard, C. (1990). Intuition in the context of discovery. *Cognitive Psychology,* 22, 72-100.

Brown, J. (2006). Contextualism and warranted assertibility manoeuvres. *Philosophical Studies*, 130, 407-435.

Casullo, A. (2003). *A Priori Justification*. Oxford: Oxford University Press.

Chomsky, N. (1988). *The Minimalist Program*. Cambridge, MA: The MIT Press.

Claxton, G. (1998). Investigating human intuition: Knowing without knowing why. *The Psychologist,* 115, 217-220.

Cummins, R. (1998). Reflections on reflective equilibrium. In M. DePaul and W. Ramsey (eds.). *Rethinking intuition: The psychology of intuition and its role in philosophical inquiry*. Lanham: Rowman and Littlefield.

Dancy, J. (1991). Intuitionism. In P. Singer (ed.), *A Companion To Ethics*. Cambridge: Blackwell Publishing.

Dancy, J. (2006). *Ethics without Principles*. Oxford: Oxford University Press.

Denes-Raj, V., & Epstein, S. (1994). Conflict between intuitive and rational processing: When people behave against their better judgment. *Journal of Personality and Social Psychology,* 665, 819-829.

Devitt, M. (2006). Intuitions in linguistics. *British Journal for the Philosophy of Science*, 57:3, 481-513.

Dorfman, J., Shames, V., & Kilstrom, J. (1996). Intuition, incubation, and insight: Implicit cognition in problem solving. In G. Underwood (ed.). *Implicit Cognition* (pp. 257-296). Oxford: Oxford University Press.

Dreyfus, H., & Dreyfus, S. (1986). *Mind over Machine: The Power of Human Intuition and Expertise in the Era of the Computer*. New York: The Free Press

Dreyfus, H., & Dreyfus, S. (1991). Towards a phenomenology of moral expertise. *Human Studies*, 14: 229-250.

Einhorn, H. J., & Hogarth, R. M. (1978). Confidence in judgment: Persistence of the illusion of validity. *Psychological Review, 85*, 395-416.

Epstein, S., Lipson, A., Holstein, C., & Huh, E. (1992). Irrational reaction to negative outcome: Evidence for two conceptual systems. *Journal of Personality and Social Psychology,* 62, 328-339.

Gendler, T. (2007). Philosophical thought experiments, intuitions, and cognitive equilibrium. *Midwest Studies in Philosophy*, 31:1, 68-89.

Goldman, A., & Pust, J. (1997). Philosophical theory and intuitional evidence. In M. DePaul and W. Ramsey (eds.). *Rethinking intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry* (pp. 179-200). Lanham: Rowman and Littlefield.

Griffin, D., & Tversky, A. (1992). The weighing of evidence and the determinants of confidence. *Cognitive Psychology,* 243, 411-435.

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review,* 108:4, 814-834.

Haidt, J., & Joseph, C. (2004). Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Daedalus*, 55-66.

Hammond, K. (1996). *Human judgment and social policy*. New York: Oxford University Press.

Hintikka, J. (1999). The emperor's new intuitions. *Journal of Philosophy* 96:3, 127 – 147

Hintikka, J. (2001). Intuitionistic logic as epistemic logic. *Synthese*, 127:1-2, 7-19.

Huemer, M. (2006). *Ethical Intuitionism*. Palgrave Macmillian.

Jackson, F. (1994). Metaphysics by possible cases. *Monist,* 771, 93-111.

Jackson, F. (1998). *From Metaphysics to Ethics: A Defence of Conceptual Analysis*. Oxford University Press: Oxford.

Kahneman, D., & Klein, G. (2009). Conditions for intuitive expertise: A failure to disagree. *American Psychologist, 64:6*, 515-526.

Kahneman, D., & Tversky, A. (1982). On the study of statistical intuitions. *Cognition,* 11, 123-141.

King, L., & Appleton, J. 1997. Intuition: A critical review of the research and rhetoric. *Journal of Advanced Nursing,* 26, 194-202.

Klienmutz, B. 1990. Why we will use our heads instead of formulas: Towards an integrative approach. *Psychological Bulletin,* 1073, 296-310.

Kornblith, H. 1999. The role of intuition in philosophical inquiry: An account with no unnatural ingredients. In M. DePaul and W. Ramsey Eds. *Rethinking intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry* (pp. 129-142). Lanham: Rowman and Littlefield.

Krosnick, J.A., & Petty, R.E. (1995). Attitude strength: An overview. In R.E. Petty & J.A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 1-24). Hillsdale, NJ: Erlbaum.

Laio, M. 2008. A defense of intuitions. *Philosophical Studies*, 140:2, 247-262.

Laughlin, C. 1997. The nature of intuition: A neuropsychological approach. In R. Davis-Floyd and P. Arvidson (eds). *Intuition: The Inside Story* (pp. 19-38). New York: Routledge.

Lehrer, K. (1990). *Theory of Knowledge*. Westview Press.

Machery, E., Mallon, R., Nichols, S. and Stich, S. 2004. Semantics, cross-cultural style. *Cognition,* 923, 1-12.

Macnamara, J. 1991. The development of moral reasoning and the foundation of geometry. *Journal for the Theory of Social Behavior,* 212, 125-150.

Miller, V. 1995. Characteristics of intuitive nurses. *Western Journal of Nursing Research,* 173, 305-316.

Monsay, E. 1999. Intuition in the development of scientific theory and practice. In R. Davis-Floyd and P. Arvidson (eds). *Intuition: The Inside Story* (pp. 103-120). New York: Routledge.

Nagel, J. 2007. Epistemic intuitions. *Philosophy Compass*, 2:6, 792-819.

Nichols, S. and Knobe, J. 2007. Moral responsibility and determinism: The cognitive science of folk intuitions. *Nous*, 41:4, 663-685.

Nichols, S., Stich, S. and Weinberg, J. 2003. Metaskepticism: Meditations in ethno-epistemology. In S. Luper (ed.). *The Skeptics* (pp. 227-247). Ashgate Publishing.

Nisbett, R.E., Peng, K., Choi, I. and Norenzayan, A. 2001. Culture and systems of thought: Holistic versus analytic cognition. *Psychological Review,* 1082, 291-310.

Osbeck, L. 1999. Conceptual problems in the development of a psychological notion of 'intution'. *Journal for the Theory of Social Behavior,* 293, 229-250

Osbeck, L. 2001. Direct apprehension and social construction: Revisiting the concept of intuition. *Journal of Theoretical and Philosophical Psychology,* 212, 118-131.

Parsons, C. 1986. Intuition in constructive mathematics. In J. Butterfield (ed.). *Language, Mind, and Logic* (pp. 211-229). Cambridge: Cambridge University Press

Parsons, C. 2000. Reason and intuition. *Synthese,* 1253, 299-315.

Petty, R., & Krosnick, J. (Eds.). (1995). *Attitude strength: Antecedents and consequences.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Plessner, H., Betsch, C. and Betsch, T. (eds.). 2007. Intuition in judgment and decision making. Mahwah, NJ: Lawrence Erlbaum.

Pust, J. 2000. *Intuitions as Evidence*. New York: Garland.

Pust, J. 2004. On explaining knowledge of necessity. *Dialectica*, 58:1, 71-87.

Redelmeier, D. and Shafir, E. 1995. Medical decision making in situations that offer multiple alternatives. *JAMA,* 273, 302-305.

Reber, A. 1989. Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General,* 1183, 219-235.

Reber, A. 1993. *Implicit learning and tacit knowledge: An essay on the cognitive unconscious.* New York: Oxford University Press.

Seung, T. 1993. *Intuition and Construction*. New Haven: Yale University Press.

Shafir, E. 1999. Philosophical intuitions and cognitive mechanisms. In M. DePaul and W. Ramsey (eds.). *Rethinking intuition: The psychology of intuition and its role in philosophical inquiry* (pp. 59-74). Lanham: Rowman and Littlefield.

Sloman, S. 1996. The empirical case for two systems of reasoning, *Psychological Bulletin,* 1191, 3-22.

Sosa, D. 2006. Scepticism about intuition. *Philosophy*, 81, 633-647.

Sosa, E. 1999. Minimal Intuition. In M. DePaul and W. Ramsey (eds.). *Rethinking intuition: The psychology of intuition and its role in philosophical inquiry* (pp. 257-270). Lanham: Rowman and Littlefield.

Sosa, E. 2000. Replies. *Nous*, 10, 38-42.

Sosa, E. 2005. A defense of intuitions. , M. Bishop and D. Murphy (eds.). *Stich and His Critics*. Blackwell Publishers.

Sosa, E. 2007. Intuitions: Their nature and epistemic efficacy. *Grazer Philosophische Studien*, 74, 51-67.

Sosa, E. 2007. Experimental philosophy and philosophical intuition. *Philosophical Studies*, 132, 99-107.

Sternberg, R.J. and Davidson, J.E. Eds. 1995. *The nature of insight*. Cambridge: Bradford/MIT Press.

Swain, S., Alexander, J. and Weinberg J. 2008. The instability of philosophical intuitions: Running hot and cold on True-Temp. *Philosophy and Phenomenological Research*, 76:1, 138-155.

Tversky A. and Kahneman, D. 1974. Judgment under uncertainty: Heuristics and biases. *Science,* 185, 1124-1130.

Ubel, P. and Loewenstein, G. 1997. The role of decision analysis in informed consent: choosing between intuition and automaticity. *Social Science and Medicine,* 445, 647-656.

Weinberg, J. 2007. How to challenge intuitions empirically without raising skepticism. *Midwest Studies in Philosophy*, 31:1, 318-343.

Weinberg, J., Nichols, S. and Stich, S. 2001. Normativity and epistemic intuitions. *Philosophical Topics,* 291-2, 429-460.

Williamson, T. 2004. Philosophical 'intuitions' and skepticism about judgment. *Dialectica*, 58, 109-53.

Wisniewski, E. 1999. The psychology of intuition. In M. DePaul and W. Ramsey (eds.). *Rethinking intuition: The psychology of intuition and its role in philosophical inquiry* (pp. 45-58). Lanham: Rowman and Littlefield.

Wright, C. 2004. Intuition, entitlement, and the epistemology of logical laws. *Dialectica*, 58:1, 155-175.

Wright, J.C., Cullum, J., & Schwab, N. (2008). The cognitive and affective dimensions of moral conviction: Implications for tolerance and interpersonal behaviors. *Personality and Social Psychology Bulletin, 34:11,* 1461-1476.

Zamzow, J., & Nichols, S. (forthcoming). Variations in ethical intuitions. *Philosophical Issues.*

Endnotes

[1] For instance, intuition has been implicated in linguistics (Chomsky, 1988; Devitt, 2006; Hintikka, 2001), rapid judgment and decision making (Griffin & Tversky, 1992; Hammond, 1996; Kahneman & Tversky, 1982; Klienmutz, 1990; Plessner, Betsch, & Betsch, 2007; Sloman, 1996), insight and problem-solving (Bowers, Farvolden, & Mermigis, 1995; Bowers, Regehr, & Balthazard, 1990; Dorfman, et al., 1996; Sternberg & Davidson, 1995), implicit learning (Reber, 1989; 1993), expertise (Dreyfus & Dreyfus, 1986, 1991), social cognition (Haidt, 2001; Osbeck, 2001; Seung, 1993), scientific theory-building (Goldman & Pust, 1998; Monsay, 1999), and even medicine (King & Appleton, 1997; Miller, 1995; Ubel & Loewenstein, 1997).

[2] It is important to note that this concern about intuition's 'epistemic' status could be targeting several things, including intuition's rational, justificatory, and/or evidential status. Thanks to [omitted for blind review] for clarifying this issue.

[3] Here I am using "stable/unstable cases" as shorthand for cases that elicit stable vs. unstable intuitions – that is, strictly speaking, it is the intuitions that are stable (or unstable), not the cases.

[4] It is important to distinguish here between conceptual *competence* and *accuracy*. While the hope is that most of the time these two will go together – that is, the competent use of our concepts will usually result in us getting things *right* – it seems nonetheless possible for them to come apart. We could envision two cultures, for instance (one whose beliefs about the nature of the universe are grounded by contemporary scientific/philosophical theory and another whose beliefs are grounded in ancient mythological lore) that might employ the concept 'knowledge' differently. Taking *Coin-Flip* as an example, while the first culture would hold that this clearly fails to count as an instance of knowledge, the latter might hold that it just as clearly counts, since Dave's "special feeling" indicates the presence of a psychic ability (or something along

those lines). While we may certainly want to say that the latter culture fails to adequately grasp the concept of knowledge (and, as such, their use of the concept in *Coin-Flip* is mistaken), we may nonetheless want to grant them conceptual competence, given that it seems reasonable to attribute knowledge to *Coin-Flip* when your belief system holds that psychic abilities (the presence of which is indicated by a "special feeling") exist. Especially since what we are interested in here is people's intuitional *stability*, not accuracy, this issue seems important to keep in mind. Indeed, such variation in underlying belief systems may help to explain the cultural variability in intuitions found by Machery, et al. (2004) and elsewhere.

[5] Zamzow & Nichols (forthcoming) found that confidence tracked instability in a set of classic ethical dilemmas (*Bystander*, *Scan*, *Transplant*) and so the inclusion of some ethical cases in Study 2 seemed prudent.

[6] Perhaps philosophical training actually expands the range of paradigmaticity – that is, through philosophical and theoretical advancement, cases that were once non-paradigmatic become paradigmatic (or cases that are non-paradigmatic for some become paradigmatic for others).

Appendix

Study 1 Cases

COIN-FLIP: 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.

TRUE-TEMP: 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.

FAKE-BARN: 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.

TESTIMONY: 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!"

Study 2 Cases

Set 1: Epistemic Vignettes

CLEAR YES (Perception): Pat walks into her kitchen during the day when the lighting was good and there was nothing interfering with her vision. She sees a red apple sitting on the counter, where she had left it after buying it at the grocery store the day before. As she leaves home, she tells her son, Joe, that there is a red apple sitting on the kitchen counter and to make sure to pack it with his lunch.

CLEAR NO (Coin-Flip): 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.

NOT CLEAR (True-Temp): Suppose Charles 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, called a tempucomp, is implanted in Charles' 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. Charles 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, at a particular moment in time he thinks and accepts that the temperature is 71 degrees – and it is, in fact, 71 degrees.

Set 2: Epistemic Vignettes

CLEAR YES (Testimony): 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!"

CLEAR NO (Guess): Laura's math teacher asks everyone to perform a difficult math problem. Laura realizes that she has no idea how to do the problem and so she just sits there and doodles. After about a minute, the math teacher asks Laura to report to the class what answer she had gotten. Not knowing what else to do, Laura blurts out "35" as a completely random guess. As it turns out, this is the correct answer and the teacher congratulates Laura for a job well done.

NOT CLEAR (Farmer): Farmer Field is concerned about his prize cow, Daisy, whom he put out into a field to graze earlier that morning. In fact, he is so concerned that he goes out to the field to check on her periodically. Standing by the gate, he sees in the distance, behind some trees, a white and black shape that he recognizes as his favorite cow. He goes back home and tells his friend, the dairyman, that he knows that Daisy is in the field, grazing happily. Yet when the dairyman leaves to go home, he walks by the field and notices that even though Daisy is in fact in the field just as Farmer Field thought, she is actually napping in a hollow, behind a bush, well out of sight of the gate (and of Farmer Field). He then also spots a large piece of black and white cardboard that has got caught in a tree, making it look like Daisy is standing there.

Set 3: Ethical Vignettes

CLEAR NO (Break-Promise): Fred promises his girlfriend that he will meet her for lunch at 12pm on Wednesday at their favorite café. Wednesday at 11:45am, on his way to the café, Fred runs into his grandfather, who is out for a stroll. They exchange hellos, and then suddenly Fred's grandfather clutches his chest and falls to the ground unconscious. An ambulance arrives minutes later to take Fred's grandfather to the hospital. Fred accompanies his grandfather to the hospital,

even though he knows that doing so means that he will be breaking his promise to have lunch with his girlfriend.

CLEAR YES (Sell-iPod): Laura and Suzy are roommates. Laura asks Suzy if she has seen her new iPod, which she had worked an extra job over the summer to be able to afford. Suzy did recently see it under a pile of papers on the bookshelf. But Suzy lies to Laura, telling her that she hasn't seen it. She thinks that if Laura doesn't find it on her own in a day or two, she can take it down to the pawn shop and get $100 for it, which would provide her with beer money for the week.

UNCLEAR (Hide-Bombers): Martha hides her Jewish neighbors in her basement during the Nazi occupation of France. A German soldier comes to her door one afternoon and asks her if she knows where her neighbors have gone. Martha knows that her neighbors are wanted by the Germans for bombing a German-only schoolyard and killing several children, injuring others. Martha lies to the soldier, telling them no, she hasn't seen them recently, but she believes that they fled the country.