Contemporary philosophical practice involves, among other things, the use of intuitional data. That is, philosophers regularly collect and consult intuitions—their own and others’—in an effort to advance, refute, and refine philosophical principles and theories. Does a particular distribution of resources seem “just”? Does a particular belief, obtained in a particular way, strike us as “knowledge”? And while it rarely makes sense to rely solely on intuitions when doing philosophy, they are nonetheless typically granted significant epistemic and dialectical weight (see, e.g., Bealer 1992, 1996, 1998; Goldman 2007; Levin 2005; Nagel 2007; Pust 2000; Sosa 2007a, 2007b) as in the well-known (Gettier 1963) paper, where the intuitions generated by a set of thought experiments were sufficient to undermine the then widely accepted account of knowledge as “justified true belief.”

This is not to say that the philosophical use of intuitions has gone unquestioned, or unchallenged. Indeed, there have long been debates about the epistemic status of intuition (e.g., whether it is legitimate, epistemically speaking, to form beliefs on their basis) driven by, among other things, disagreements about the nature of intuition (Audi 2004; Bealer 1999, 2000; Bengson 2014; Chudnoff 2011; Claxton 1998; Huemer 2006; Kornblith 1998; Laughlin 1997; Osbeck 1999, 2001; Parsons 2000; Pust 2000; Sosa 1998, 2007a, 2007b; Williamson 2007; Wisniewski 1999) and the cognitive processes intuiting might be or involve (Cummins 1998; Denes-Raj and Epstein 1994; Dorfman, Shames, and Kilstrom 1996; Koksvik 2013; Epstein et al. 1992; Gendler 2007; Osbeck 1999; Shafir 1999; Sloman 1996).

The past few decades have seen an increase in these questions and challenges (see, e.g., Bishop and Trout 2005; Cummins 1998; Denes-Raj and Epstein 1994; Devitt 1994; Elgin 1996; Gendler 2007; Hintikka 1999, 2001; Kornblith 1998; Machery et al. 2004; Nichols and Knobe 2007; Nichols, Stich, and Weinberg 2003; Nisbett, et al. 2001; Redelmeier and Shafir 1995; Stich 1988; Weinberg 2007; Weinberg, Nichols, and Stich 2001), generating serious concern about intuition’s value in philosophical practice. One such challenge comes in the form of empirical
research suggesting that people’s “intuitional judgments” (i.e., judgments based upon, driven by, or the result of an intuition) about philosophically-relevant cases are vulnerable to bias. Specifically, they have been found to be *inappropriately sensitive* to things – for example, “aspects of who we are, what we are being asked to do, and how we are being asked to do it” (Alexander and Weinberg, 2014) – that have nothing to do with the cases being considered and/or the truth or falsity of the claims being made. This results in problematic (and according to some, unresolvable) variance or “instability” in people’s intuitional judgments.

The studies documenting this instability are too numerous to list in their entirety here, but to follow are some highlights:

- Nichols and Knobe (2007) found that people’s judgments were influenced by whether they were considering a concrete or abstract case, presumably because the former triggered a “biasing” emotional response.

- Petrinovich and O’Neill (1996) found that people’s judgments were strongly influenced by “framing effects” (Tversky and Kahneman 1981), specifically by whether they were encouraged to focus on the number of people who *would be saved* or the number of people *who would die* because of their chosen action – the numbers being the same across both cases.

- Swain, Alexander, and Weinberg (2008) found that people’s responses to concrete cases were vulnerable to an “order effect” (Tversky and Kahneman 1974), their judgments significantly influenced by the case they had previously considered (see also Liao, Wiegmann, Alexander, and Vong forthcoming; Nichols and Zamzow 2009; Petrinovich and O’Neill 1996). And other research suggests this instability is not simply an artifact of shallow reflection – Weinberg, Alexander, Gonnerman, and Reuter (2012) found order effects in the judgments of people dispositionally inclined towards high levels of cognition (though interestingly in the opposite direction) and Schwitzgebel and Cushman (2011) found order effects in philosophers themselves.

There have also been studies suggesting that people’s judgments vary as a function of differences in their cultural backgrounds, socioeconomic status (Weinberg, Nichols, and Stich 2001; see also Machery et al. 2004) and gender (Buckwalter and Stich 2011; see also Nichols and Zamzow 2009; Petrinovich, O’Neill, and Jorgenson 1993). But, it is important to be clear that these differences likely represent a different sort of worry from the problem of other cognitive biases (such as framing and order effects). The latter involves intuitive judgments being unduly influenced by information present/salient *at the moment our judgments are formed*, while the former involves a much more complex story about the ways in which sociocultural belief systems/norms become internalized, shaping our understanding and use of certain concepts – and perhaps even the concepts themselves. Though such differences may be at the heart of some interesting cultural differences in our philosophical theories, we would not typically regard them as a form of “bias” (at least not in the way we regard the latter to be). 5

These findings, taken together, bring into question philosophers’ reliance on intuitions as sources of evidence/justification in philosophical practice. Or so goes what has come to be known as the “restrictionist challenge”, which maintains that the instability found in people’s intuitional judgments represents “a worrisome methodological deficiency in philosophers’ armchair practice of appeal to intuitions” (Weinberg et al. 2012, 257; see also Alexander and Weinberg 2007) and, more, “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” (Swain, Alexander, and Weinberg 2008, 1).

Yet, even in the wake of these findings – and the challenge to philosophical methodology they are taken to represent – there are many who continue to defend the use of intuitions as a
reasonable (even “standard”) part of philosophical practice (e.g., Bealer 1998; Cullen 2010; Deutsch 2009, 2011; Devitt 2012; Kauppinen 2007; Ludwig 2007; Moffett 2007; Sosa 2005, 2007a; D. Sosa 2006; Williamson 2009). One important line of defense given is that we cannot draw conclusions about the epistemic status of intuitions from these studies, as conducted, because “[i]t cannot innocently be assumed that subjects’ answers expressed how things struck them – what intuitions they had, if any” (Bengson 2013, 496). That is, we cannot simply assume that the participants in these studies were forming intuitional judgments, because it is just as (if not more) likely that they were doing something else entirely – for example, guessing, giving responses that they deemed socially suitable/acceptable, and so on – and none of the studies conducted thus far have attempted (much less successfully managed) to control for this.6

Relatedly, Ludwig (2007) argues that it would be very difficult to assess whether – and, thus, irresponsible to assume that – participants’ responses were the sorts of judgments we are actually after, namely, “judgments which express solely the subject’s competence in the deployment of the concepts involved” (italics mine, 144–145). After all, there are so many other, as yet unmeasured, factors having little or nothing to do with people’s conceptual competence – performance errors, pragmatic concerns, social desirability, and so on – that could have been influencing/informing their judgments (see also Cullen 2010; Kauppinen 2007).

In short, this line of defense calls for a distinction between people’s intuitions and their judgments. And while we can clearly agree that the studies in question successfully documented an instability in people’s judgments, they did not successfully demonstrate that the source of that instability is intuitional in nature – that is, that it is people’s intuitions that are being “pushed around.”

If taken seriously, these rejoinders to the restrictionist challenge essentially send us back to the drawing board. The question of whether we have good reason to believe that people’s intuitions are problematically unstable remains open until someone solves the logistical problem of verifying that what we are measuring is people’s intuitions – in the absence of, or controlling for, other potentially distracting/impeding factors – and not some other mental states.7

But there is another line of defense worth noting because of its relevance even if we assume that the studies in question actually did (more often than not) capture people’s intuitional judgments. For, as Laio (2008) points out, even if we take seriously the findings that some intuitional judgments were influenced by cultural background, socioeconomic status, order of presentation, and so on, we must also then take seriously the fact that others were not (see also Petrinovich and O’Neill 1996). One of the largely unacknowledged gems of the studies in question is that, in the midst of all the instability discovered, there was stability as well. In short, the evidence suggests that while some intuitional judgments (assuming they are intuitional) are vulnerable to a variety of problematic biasing influences, others may not be.

Of course, it is not immediately clear how much defensive traction can be gained from this observation. There are at least two concerns with this line of defense that deserve consideration. The first is that the existence of stable intuitional judgments does not buy us much because, as a group, they are philosophically “uninteresting” (Laio 2008).8 And, second, even if they were not entirely dismissible on these grounds, there is still the worry that we have no way of knowing (short of conducting additional studies) which intuitional judgments will be stable and which will not (see also Alexander and Weinberg 2007). In what follows, I will address the latter of these two worries first, leaving a discussion of the former until the end.

One response given to this latter concern (Sosa 1998; Williamson 2004) is that discovering instances of instability in our intuitional judgments does not thereby restrict us to having to confirm each time, or be able to anticipate ahead of time, which of such judgments will be stable. After all, our awareness of the potential unreliability of judgments we form on the basis of perception and memory does not force us to have to continuously verify that a given perceptual
or memory-driven judgment is reliable – we are typically justified in relying on them without such confirmation.

While this certainly seems right, it is not the only response available to us. There is evidence to suggest that people may be (at least indirectly) aware of when their intuitional judgments are stable. This evidence comes from research I have conducted to investigate intuitional instability (Wright 2010, 2013), which resulted in two discoveries:

1. Across multiple studies there was a subset of stable cases (i.e., cases that elicited stable intuitional judgments) – for example, in Wright (2010) two-thirds (6 of 9) of the epistemological and ethical cases presented generated judgments that were stable across order manipulations.

2. People successfully “tracked” this stability, in the sense that their confidence in their judgments, and the strength with which they believed their content, predicted judgment stability. People reported being significantly more confident in, and believing more strongly, the judgments that were stable against manipulation.

What makes an intuitional judgment stable? My suggestion is that stable intuitional judgments are those judgments that involve (i.e., are based upon, driven by, or the result of) clear/strong intuitions. When someone has a clear/strong intuition about a case (e.g., the intuition that “x is p”) – and no overriding reasons to ignore/discount those intuitions – they will tend to form judgments on the basis of that intuition (i.e., the judgment that “x is p”). They will also tend to feel confident in those judgments and to believe strongly their content (i.e., that x is p) – after all, it strikes them clearly and strongly that it is the case that x is p. Since these attitudinal states (confidence and belief strength) are states for which people have “introspective access,” they can thereby serve as reliable indicators of stability (see also, Nichols and Zamzow 2009).10

In order to further test the relationship between stability and confidence/belief strength, I induced instability in people’s previously stable judgments by introducing information designed to interfere with, or cause people to override, their clear/strong intuitions – in this case, expert disagreement that was either consistent or inconsistent with people’s stable judgments – and found by doing so resulted in a corresponding decrease in people’s confidence in their judgments and strength of belief of their content. That is, not only did their judgments become unstable, but this instability was accompanied by a significant decrease in confidence/belief strength (Wright 2013).

Crucially, this decrease in confidence/belief strength in their now destabilized judgments was not simply the result of a decrease in certainty brought on by exposure to expert disagreement about which was the correct judgment to form. If a link between confidence and certainty was driving the effect, then we would expect the presence of expert consensus (which, arguably, should increase certainty in one’s judgments) to correspond with increased confidence/belief strength. Yet, a follow-up study revealed that even in the presence of expert consensus, people’s judgments about unstable cases were still accompanied by reduced confidence and belief strength – even when their judgments aligned with that consensus.

For example, even when people read of True Temp Charles that over 100 professional epistemologists and linguists agreed that he knew (or did not know) the temperature, their judgments were still manipulated by the order effect – this time by aligning with the expert testimony – and their confidence and belief strength were still significantly lower than for any of the stable cases that preceded and followed. Thus, instability was associated with reduced confidence/belief strength even when, if anything, people had reason to feel more confident about their judgments, not less.

Taken together, these findings suggest that people’s introspective access to certain attitudinal states serves as a reliable indicator of intuitional stability: the level of confidence they experience
for their judgments, and strength with which they find themselves believing their content, predict whether or not their judgments will display stability across manipulations. And they most likely predict stability (at least in the case of intuitional stability) because they reflect striking the judger – clearly and strongly – that x (whatever x may be) is (or is not) the case.

This would explain why the confidence/belief strength people reported for True Temp was low, despite largely aligning with expert consensus – despite being told that the experts agreed about whether True Temp knew (or didn’t know) the temperature, it just didn’t strike them one way or the other. And so, they formed their judgments on the basis of the expert testimony – not on the basis of a clear/strong intuition.

And it further suggests that one reason for unstable judgments is the lack of clear/strong intuitions about particular cases (having only weak/unclear intuitions or no intuitions at all) – in the absence of which people must rely on other information (e.g., the previous case considered, expert testimony, etc.) to inform their judgments.

Another reason for unstable judgments is that, even in the presence of clear/strong intuitions, other things (e.g., conflicting information, strong emotions, social pressures, etc.) can get in the way, causing people to question/ignore/override them – intuitions are, after all, only one of many potential sources of information on the basis of which judgments can be formed. But, however the instability is generated, there are certain assessable attitudinal states (viz., confidence/belief strength) that can serve to reliably predict its presence.

Of course, as I’ve already noted, one of the upshots of this is that intuitional instability might not be intuitional at all – which means it may not be people’s intuitional judgments, per se, that we should be concerned about, but rather the judgments they form “under the influence” of so many other potentially biasing factors. Specifically, we should be concerned about the judgments people form in the absence of clear/strong intuitions – where, in seeking elsewhere for information to determine the appropriate response, they become vulnerable to biases that take a myriad of forms. As Bengson (2013) notes, we have good reason to doubt the epistemic status of people’s judgments when they are the result of things like “quick hypotheses, obliged guesses, hasty inferences, and hurried emotional reactions” as these “are not legitimate epistemic sources. Arguably, these mental states are among those that [we] can mutually agree have no positive epistemic status.” (521)

Likewise, we should be concerned about the various things – for example, social desirability, peer pressure, expert testimony, prior theoretical commitments, and so on – that can throw people off their clear/strong intuitions, resulting in judgments that contradict with how things strike them. The objective in both cases is to generate environments that maximize our sensitivity to the presence/absence of clear/strong intuitions and minimize the sort of interferences likely to derail them. And this sort of vigilance should be reasonably easy to incorporate into standard philosophical practice, to the extent that it isn’t already.

This brings us back, however, to the other concern about this line of defense, which questions the philosophical value of the stability in question. Elsewhere (Wright 2010, 2013), I’ve hypothesized that one important factor in the generation of clear/strong intuitions is the “paradigmaticity” of the cases under consideration – that is, the degree to which they represent clear instances of the concepts in question, and therefore elicit clear/strong intuitions. For example, when considering whether or not someone “knows” something, some cases provide clear instances of knowledge (e.g., beliefs gained through direct perceptual observation under ideal conditions), while others provide clear instances of not knowledge (e.g., randomly guessing the answer to a question, even if you happen to guess correctly). Cases such as these are clear – and they strike us as such. And, as some (e.g., Goldman 2007; Sosa 2000) have argued, such “strikings” have probative force – we feel their pull and are “attracted to assent” to their content. Thus, when considering such cases, our judgments are likely to be more stable – much less likely to be influenced.
by outside factors — than when we are considering difficult, borderline, or highly complex, cases (cases that may strike us in multiple ways or fail to strike us at all).

But, herein lies the rub (as they say), because important philosophical work often goes on “at the margins,” involving complex concoctions of thought experiments that push beyond of our conceptual and experiential comfort zones. This observation not only renders paradigmatic cases — and the clear/strong intuitions they generate — philosophically uninteresting, but it leaves the philosopher with cold comfort, because it is precisely were instability is most likely to lurk that she may need to rely the most heavily on her intuitions.

This is certainly an important consideration, to which I have only a couple of (inadequate) thoughts. The first is that calling the stable cases “philosophically uninteresting” fails to give sufficient weight to the role of paradigmaticity, which arguably provides the conceptual “bedrock” necessary for philosophical development to occur. Identifying the boundaries of our concepts — distinguishing the sorts of cases that strike us as clear instances from those that don’t — is arguably a valuable endeavor in its own right (see also Laio 2008).

Secondly, when it comes to clear/strong intuitions, philosophers are clearly not restricted to paradigmatic cases — they can have clear/strong intuitions about other kinds of cases as well. Such cases, while potentially confusing and therefore vulnerable to bias for the general population, may nonetheless elicit stable intuitional judgments for philosophers who have received extensive training designed to refine and enhance their conceptual mastery (at least in their area of interest) and their capacity to see various inferential/logical connections. Such training arguably provides a greater capacity for discrimination and, therefore, the ability to “intuit” difficult cases more clearly, and more difficult cases clearly, than the novice.

What is more, the very act of philosophical discourse itself generates new levels conceptual clarity, expanding and refining our range of conceptual competence and our ability to see various connections (entailments, etc.) between them. Perhaps philosophical training also changes (through conceptual and even inferential connections) the scope of paradigmaticity, expanding the range of cases that are recognized as clear instances of a concept through philosophical and theoretical advancement.

In conclusion, when it comes to relying on concrete case judgments for philosophical advancement, one thing this debate has helped to make salient is the importance of generally attending to our intuitional judgments, being sensitive to whether they were formed solely on the basis of clear/strong intuitions — and being wary of cases for which our intuitions are “fuzzy” or absent — as well as vigilant against the sorts of cognitive and situational influences that can throw off our judgments, even when a clear/strong intuition is present.

Although it may turn out that the many studies intended to document instability are not, in the end, critiques of intuition, we can nonetheless agree that they represent an important (though not surprising) cautionary tale about the many challenges we face when it comes to making good judgments.

Notes

1 I take this claim to be relatively uncontroversial. However, it is worth nothing that it has recently been challenged in Cappelen’s book (2012). Despite his arguments that “methodological rationalism” should be abandoned — and experimental philosophy along with it — I am going to stubbornly stand by my claim (otherwise, what would be the point of this chapter?) without defending it. But for those interested in a defense, there are some excellent replies to Cappelen forthcoming in a special edition of Philosophical Studies.

2 Moving forward, let us assume a relatively thin notion of intuitions — namely, that they are (or involve) the experience of something seeming to be, or striking you as, the case — “when one has the intuition that
it seems to one that \( p \) (Koksvik, 2013, 3); or, as Chalmers (2014) put it, “intuitive claims seem obviously [i.e., not requiring further “broadly inferential” justification] true” (3).

See, for example, Weinberg (2007).


Take Coin-Flip Dave (Swain, Alexander and Weinberg 2008) as an example: While one culture might hold that this case clearly fails to count as an instance of knowledge, another might hold that it just as clearly counts, since Dave’s “special feeling” indicates the presence of a psychic ability – something this culture clearly believes to be a reliable source of information (at least for those who possess it). Thus, while we might want to argue that the latter culture holds false beliefs about the nature of psychic abilities or even fails to adequately grasp the concept of knowledge (and, as such, their intuitions about Coin-Flip Dave are mistaken), we would not want to say that their intuitions are biased, given that it seems reasonable to attribute knowledge to Coin-Flip Dave when your belief system holds that reliable psychic abilities (the presence of which is indicated by a “special feeling”) exist.

As Bengson further points out, this worry does not hinge upon any sort of “elitist” (or otherwise restrictive) account of intuition or appeal to some distinct “sub-species” of intuitions. Rather, it follows straightforwardly from the following two uncontroversial observations:

1. “Sometimes things strike us a certain way; other times they do not, even though we may still answer a question about whether things are that way when prompted.” (508)
2. “Sometimes things strike us a certain way, even though we may answer that things are not that way, but rather some other way, when prompted.” (511)

Of course, the “restrictionists” have not taken this challenge to their challenge lying down – e.g., see Weinberg and Alexander (forthcoming), as well as an interesting discussion in Sytsma and Livengood (forthcoming) – specifically Ch 4:2.

Consider: People overwhelmingly judge that “Pat knows that there is an apple on the table” when she has direct visual perception of an apple on the table in front of her and no reason to doubt that perception – and their judgments in this regard are stable against the manipulations discussed above. But, there is not likely to be much in the way of (new) philosophical insight into the nature of knowledge to be had by this fact, or by the case itself (a version of which – called, aptly, Perception – was used in Wright 2010, 2013).

Although it is not without its critics—see, for example, Weinberg (2007) and Alexander (2012) for arguments against this defense.

This is not to say that there aren’t other kinds of stable judgments, or other reasons why our judgments might be stable. We may, for example, have a stable judgment about something simply because we stubbornly refuse to change our judgment, even in the face of conflicting evidence (including conflicting intuitions). I take it, though, that these sorts of judgments aren’t intuitional.

It is also important here to be clear that what we are talking about here is stability, not accuracy. That someone has a clear/strong intuition about a particular case obviously does not guarantee that her intuitive judgment will be correct, even granting her reasonable conceptual competence. Other beliefs she holds or theoretical commitments she has may influence (even distort) her grasping of particular concepts, thereby influencing which cases clearly strike her and which don’t. For some interesting research on the relationship between confidence and accuracy, see Koriat (2008).

For example, in cases like Perception (discussed in Footnote 8).

In the True Temp case, Charles has the ability to reliably tell the current temperature (for reasons that vary across different versions of the case, from a surgical implant to being hit on the head) but has no idea that he can do this or why—he just periodically forms beliefs about the current temperature (that happen to be true).

Notice how these reasons for instability hearken back to Bengson’s (2013) two observations (footnote 6).

Of course, for people (e.g., philosophers) trained to be sensitive to certain qualities of their intuitions—such as their clarity and strength—a further reference to the degree of confidence they experience in their intuitive judgments or the strength of their belief of their content may not be necessary. But for the layperson, confidence/belief strength can serve as a reliable “proxy.”

Examples that come to mind for me are cases like Davidson’s Swampman or Chalmer’s Zombie cases, though I’m sure there are other, better, examples out there.
INTUTIONAL STABILITY

16 This sort of refinement through training is not uncommon—consider, for example, being taught how to recognize the difference between a *Spizella passerina* (Chipping Sparrow) and a *Passer domesticus* (House Sparrow), which thereafter gives one the ability to distinguish between two birds that before that seemed indistinguishable. For skepticism about this view, however, see Alexander’s chapter (this volume).

17 At least one philosopher that I’m aware of argues that intuition can result from an inferential reasoning (Koksvik 2013).

References


JENNIFER COLE WRIGHT


INTUITIONAL STABILITY


