

January 2012

Intuition & Calibration

Jonathan M. Weinberg
University of Arizona

Stephen Crowley
Boise State University

Chad Gonnerman
Indiana University

Ian Vandewalker
Center for Reproductive Rights

Stacey Swain

Follow this and additional works at: <https://commons.pacificu.edu/eip>

 Part of the [Philosophy Commons](#)

Recommended Citation

Weinberg, Jonathan M.; Crowley, Stephen; Gonnerman, Chad; Vandewalker, Ian; and Swain, Stacey (2012) "Intuition & Calibration," *Essays in Philosophy*: Vol. 13: Iss. 1, Article 15.

Intuition & Calibration

Jonathan M. Weinberg, University of Arizona
Stephen Crowley, Boise State University
Chad Gonnerman, Indiana University
Ian Vandewalker, Center for Reproductive Rights
Stacey Swain

Published online: 30 January 2012

© Jonathan M. Weinberg, Stephen Crowley, Chad Gonnerman, Ian Vandewalker, Stacey Swain 2012

Abstract

The practice of appealing to intuitive judgments concerning esoteric cases, long standard in analytic philosophy, has recently fallen on hard times. Various recent empirical results have suggested that philosophers are not currently able to distinguish good intuitions from bad. This paper evaluates one possible type of approach to this problematic methodological situation: calibration. Both critiquing and building on an argument from Robert Cummins, the paper explores what possible avenues may exist for the calibration of philosophical intuitions. It is argued that no good options are *currently* available, but leaves open the real possibility of such a calibration in the future.

1. The Restrictionist Challenge for Philosophical Intuitions

Intuitions and their role in philosophical methodology have been a matter of much discussion in recent years. Discussions that have grown out of the dissatisfaction felt by some philosophers with the adequacy of intuitions to play the evidential role apparently assigned to them by many philosophers. The nature of these worries are various but it may fairly be said that they take on an empirical flavor with Steve Stich's work in the late 80's (See Stich (1988); and DePaul and Ramsey (1998) for a nice summary of this early period, both pro and con intuitions), a flavor enriched in a variety of ways by subsequent work (e.g., Bishop & Trout (2005)). Much of the debate in the last decade has turned on the results of experimental philosophy studies¹, whose relevance and impact, and indeed whose very standing as philosophy, is still a matter of intense debate.

We will follow the recent usage of Alexander & Weinberg (2007) and call this set of

Corresponding Author: J. M. Weinberg
University of Arizona
email - jmweinberg@email.arizona.edu

empirical arguments aimed to revise or abandon our current armchair practices *the restrictionist challenge*. In brief, the challenge has three parts. First, there are the experimental philosophers' empirical results themselves, that purport to reveal particular patterns of responses in ordinary subjects (typically, but not exclusively, university undergraduates). Second, there is a metaphilosophical premise that the pattern revealed in those experiments is not one that is well-aligned with the relevant philosophical truths. For example, intuitions about reference seem to vary with culture, but the fundamental facts about reference perhaps ought not so vary (Machery et al. (2004)). Third, there is an abductive inference from the observed patterns in the folk's intuitions to the same patterns afflicting the intuitions as deployed by philosophers in their armchair practice. Taken together, these underwrite an inference to the claim that there is a worrisome methodological deficiency in philosophers' armchair practice of appeal to intuitions.

There are numerous strategies of response available to the would-be defenders of the armchair, who we will term *cathedrists*. They can try to undercut the empirical results themselves, perhaps by offering deflationary explanations of them. They can reject the metaphilosophical claims. For example, they could propose a form of contextualism that can accommodate the order effects observed by Swain *et al.* (though see the discussion of that issue in that paper). They can try to block the inference from the folk to the philosophers, perhaps by conjecturing that philosophers have some special expertise that will render them sufficiently immune from these patterns² or that they have further norms of theory-selection that will overcome these sources of error.³ These strategies all attempt to *block* the restrictionist challenge: if they are successful, then they would show that philosophers' appeals to intuition only *prima facie* appears to be challenged, but in fact on closer investigation does not face such a challenge yet.

The prospects of such strategies ought to be explored, and many already are. But certainly none have already been successful, and our interest in this paper concerns the question: what if no such strategies are successful, and the restrictionist challenge cannot be blocked? If the challenge cannot be *blocked*, can it still perhaps be *met*?

Contemporary analytic philosophy should not feel strangely singled out in finding itself potentially facing such a question. For the history of inquiry is replete with examples of instruments and practices that have faced challenges to their reliability, which were then overcome by making suitable adjustments and changes to them.

For example, scientists in the early 20th century began to become aware of the unfortunate fact that human observation could be badly colored by the hopes, desires, and expectations of the observer. We will not attempt here anything like a comprehensive history of this phenomenon, though we suspect that any such story would start somewhere around the saga of Robert Wood and René Blondlot's nonexistent "N-rays" in the world of *fin-de-siècle* physics. The most fundamental fix for this methodological

flaw reached its first full fruition in 1937, when Gold et al. published a paper on the treatment of cardiac pain in which appears the first documented appearance of “blind” as a methodological term (Strong 1999), and in which both subjects and scientists were kept in the dark as to which subjects were in which experimental groups. Science triumphed not by refusing to change in response to a documented problem, but by devising an elegant solution to that problem, with a methodological change that was rapidly promulgated and adopted as a new experimental norm. In the terms of the restrictionist challenge, they faced a divergence between their patterns of judgment and how they thought the relevant truths should go, and they adopted a change to their practices so that judgment and fact would no longer diverge in that way.

The sort continual tweaking of an instrument in pursuit of improved performance that we have in mind is also well illustrated by the history of telescope. Early telescopes were objects of wonder but also subject to serious flaws. The lenses of these telescopes generated serious problems. For example, because light of different wavelengths refracts at different rates the different components of white light were focused in slightly different places yielding images with colored fringes.⁴ Initially astronomers responded to this issue (and others) by reducing the curvature of the lens – this does not eliminate the fringes but makes them much less noticeable. Subsequently, beginning with the work of Chester Moore Hall in the 1720’s and continuing into the 19th century telescope makers manufactured compound lenses whose parts were composed of glass with differing refractive properties.⁵ These compound lenses were able to focus light of different wavelengths at the same point, so that finally all the various wavelengths could be focused at the same point. Thus, over the course of nearly 300 years users and manufacturers of telescopes co-operated to improve the performance of their instrument of choice to the point where this particular problem was for all practical purposes fully resolved.⁶ We would conjecture that this cycle of becoming aware of a source of error, and then taking active steps to eliminate it, plays a large part in the success of the sciences on the whole over the last few centuries.

Analytic philosophers are perhaps not used to contemplating methodological revisions and tweaks, as our epistemological questions tend to be pitched at a grander scale, such as: How is knowledge of other minds possible *at all*? Is rational intuition a sense-independent source of justification? Does justification completely supervene on the internal psychological states of an agent? While patently legitimate questions, they nonetheless happen to operate at a height far removed from methodological particulars. Moreover, the areas of inquiry and the history of inquiry that have caught philosophers’ eyes tend also to be either revolutionary in scope (as with modern physics, generative linguistics, or the birth of analytic philosophy itself), or without any obvious resolution (as with Bayesian vs. frequentist statistics, or dynamicist vs. classical approaches to cognitive science). Alterations to practice which have proved obviously, boringly successful tend to fly beneath our professional radar.

Although we have given primarily scientific examples, we do not think there is anything *recherche* or highfalutin in the idea of revising practices that have been revealed to have some epistemic flaw, even when that flaw has proved non-lethal. It seems to us to be simply the unspoken corollary to “if it ain’t broke, don’t fix it” – namely, “wherever and to what extent *is* broke, *do* fix it.” And note that *fixing* need not be the same as *discarding*. George Bealer has appealed to our “standard justificatory practices”, and the role of intuition in them, as a way of defending intuition from would-be radical empiricists who aim to expunge intuitions altogether from our epistemic repertoire. “It is standard justificatory practice,” he writes, “to use intuitions evidentially. Unless and until a reason for departing from this standard practice is produced, we are entitled – indeed, obligated – to continue using intuitions as evidence” (1996, p. 30, n15). But it is *also* part of that standard practice to keep an eye out for exactly those sorts of reasons to depart from the practice – that is, the practice in its *current* form. Yesterday’s standard practice is today’s textbook example of previously-unforeseen methodological differences; and *mutatis mutandis* for today’s standard practices and tomorrow’s textbooks.

So we should be interested, as philosophers, in the prospects for revisionary-even-if-not-revolutionary approaches here. If the restrictionist challenge cannot be blocked outright, then should any such approaches be successful in helping at least partially meet that challenge, our practices will be better for it. And should none of them succeed, then, knowing that, we can more squarely face the possibility of meeting the challenge by jettisoning the practice in question. In this paper, we will articulate the possibility of, and evaluate the current prospects for, a particular variety of the more moderate kind of response. We suspect that there may be more than one such moderate response that would be possible, but we will focus here on what we take to be the most obvious candidate: calibration. We will begin by offering a brief and preliminary account of what calibration is, in order to set up our presentation and critical evaluation of an important recent argument from Robert Cummins for the impossibility of any epistemically valuable calibration of philosophical intuitions. In that evaluation, we will contend that Cummins is operating with too impoverished a notion of what can enable a successful calibration. As a result he ignores a variety of calibrational resources that are, at least in principle, deployable. But in-principle possibility of course fails to entail actual, current plausibility, and our final task will be to canvass the prospects for philosophers in the early 21st century wishing to pursue this kind of revision to their intuitive practices. We will conclude with a cautious and limited optimism, albeit one that proponents of a pure and autonomous armchair for the philosopher will not find satisfying.

2. The process of calibration

Suppose you have a putative source of evidence in some domain, but for which you are concerned that it may have some nontrivial way in which its deliverances diverge from the facts in that domain. We take it that the most standard way to restore trust in such an impeached source is to *calibrate* it.⁷ Cathedrists might plausibly hope that this process, so valuable in the sciences, can rescue their preferred philosophical methodology from an ignominious end. Our project here is to evaluate the prospects for such a calibration, and to do so, we will begin with a general though by no means complete picture of calibration.

Calibration is a process of regulating a putative source of evidence, by inspecting it and, if needed, adjusting it to render it accurate. Calibration is divisible into three parts: *testing*, *diagnosis*, and *correction*.

Testing is a matter of checking the accuracy of the target device over the intended domain.⁸ During testing, we use the device to make a series of observations of phenomena within the device's intended domain of application.⁹ We then attempt to evaluate the accuracy of those observations, typically by checking them against what we have reason to expect based on the deliverances of other sources, or of our understanding of the general contour of facts in the domain, or both. For example, a scientist might determine that a mercury thermometer is faulty if its readings do not compare well with those of an already trusted platinum thermometer; or just by applying it to a sample as it is heated, which can be expected to increase the temperature at least somewhat, even without our knowing the exact amount.¹⁰ If it gives a different result from the trusted device, or fails to register an increase in temperature as it heats up, then it may count as failing the test. This latter sort of test is important, as it allows a device to be shown problematic over a domain even without our knowing what the right answers in that domain.

If the process of testing reveals that the target device is sufficiently accurate, then there is no need to proceed to the other two stages of calibration. On the other hand, if an undesirable amount of deviation is found – and if we want to continue using the device – then the next step in a calibration will be to determine in what way the device goes awry. Optimally, we would have a worked-out theory of the source of error; for example, physics enables us to understand the circumstances in which a pocket compass will fail to point north, such as when it is near a large iron deposit or a coiled electrical wire. Diagnosis may still go forward in the absence of such a theory, to whatever extent we can tell at least in what kinds of cases the device tends to be more reliable, and where less.

If testing has revealed an undesirable amount of deviation in our device, and we have a diagnosis of this deviation, then we finally face the question: how do we go about correcting for this deviation? There are at least two methods for correction: *restriction*

and *rehabilitation*. With restriction, as the name suggests, we simply ignore the device's results in circumstances when we expect it to be unreliable.¹¹ With rehabilitation, we tweak the problematic device itself, rendering it accurate across its entire domain.¹² In order to make these two methods clearer, we will consider their application in a case borrowed from science: the astronomer's problem of seeing.

'Seeing' is a technical term in astronomy, referring to the blurring and twinkling of an image caused by turbulence in the Earth's atmosphere.¹³ Albeit romantic in nonscientific contexts, this phenomenon is a menace to astronomers. Although we have good astrophysical reasons to suspect that celestial bodies are themselves neither blurry nor twinkly, our naïve telescopic observations would lead us to believe that they are both. The testing stage, thus, indicates something amiss with the deliverances of the device as things stand.

We can turn to optics to diagnose the source of these errors. Each time light from a celestial object hits a pocket of air with an index of refraction that is different from its surroundings, the light from the celestial object is bent, ultimately producing a blurred image as well as presenting astronomers with the frustrating fact that they rarely acquire images as sharp as the diameter of their telescopes allow. Astronomers can correct for the undesirable effects of seeing via both restriction and rehabilitation. First, astronomers can *restrict* their usage of the telescope to circumstances within which it will produce sharp, untwinkling images – for example, by observing only large and proximate objects, like the moon. Astronomers dissatisfied with such a restriction can correct for the effects of seeing by *rehabilitating* the telescope itself, for example, by changing it to one that uses adaptive optics and flexible mirrors, or, as with the Hubble Space Telescope, by relocating the telescope outside the atmosphere altogether.

So, if the restrictionist challenge cannot be fended off and must instead be met, then calibration should seem like a very tempting way to do so. It has a long and highly successful track record in the sciences, and indeed those successes were often with similar sorts of worries put forward in the restrictionist challenge. And, moreover, one might think there would be no theoretical reason why it should not be a possible strategy for a friend of intuition to pursue.

At least, one might think that, until reading Rob Cummins.

3. Cummins vs. Calibration: The Problem of Novelty & The Prospects for Extrapolative Calibration

In his (1998), Cummins offers a powerful argument against the calibration of intuition. It seems that intuitions can be calibrated only if we *already* have trusted non-intuitive

access to the facts of its intended domain of application. And we have this sort of access, it seems, only when we have a trusted, non-intuitive theory covering that domain. But with this sort of theory in hand, there is no epistemic work left for the intuitions to do! Therefore, Cummins concludes, calibration of philosophical intuition is possible only when it would be useless:

philosophers could have no possible use for intuitions in a context in which the relevant theory was well enough settled to form the basis of a credible calibrated test. Philosophical theory in such good shape is ready to bid the Socratic midwife farewell and strike out on its own in some other department. Philosophical intuition, therefore, is epistemologically useless, since it can be calibrated only when it is not needed (Cummins 1998: 117-8).

Cummins' argument is somewhat compressed, and in particular, he does not fully articulate the required premise that *any* independent theory sufficient to calibrate intuition on the one hand, would thereby render those intuitions useless on the other. Although we think that this premise is, in fact, false, we take it to be quite suggestive. In the remainder of this section, we will try to clarify what we take to be a basic plausible connection between the required existence of independent evidence, and the possible usefulness of the calibrated intuitions. We will term this connection the *novelty requirement*. Once we see how the novelty requirement works, we will be able to see how Cummins is too quick in presuming that the calibration of intuitions will violate it, and once that error is recognized, it will open up the possibility of sources that could be the basis for a genuine and epistemically novel calibration of intuitions.

The novelty requirement states that, in order for a calibration to be epistemically useful, the deliverances of the calibrated device must be able to *extend* our knowledge to hitherto unevaluated propositions. Such novelty thus requires that one's trusted independent access to the facts of the domain in question cannot cover the entire domain of application of the device being tested. If it did so cover, then there would be nothing to be learned from the device, no matter how well-calibrated it is, and thus nothing that the device can teach us that we could not already know through other means.¹⁴

On Cummins' view, there must be some subset of the device's domain for which we have trusted independent access, in order for calibration to even occur. The subset forms a certified basis for calibration, from which one can check and, if necessary, correct the device's reliability over its entire domain. As we read him, Cummins thinks that when the certified basis comes from a trusted, independent theory, there is no room left for an expansion of our epistemic reach: the certified basis will cover the entirety of the domain of the device being calibrated. As he claims, "philosophers could have no possible use for intuitions in a context in which the relevant theory *was well enough settled to form the*

basis of a credible calibrated test.” (Cummins 1998: 118; emphasis added). Put in our terms, he claims that possession of a certified basis from a trusted, independent theory will *always* entail a violation of the novelty requirement. To reach this conclusion, Cummins assumes that an independent theory capable of providing a certified basis for calibration will also thereby provide the resources to answer all questions within the relevant domain. He offers no argument for this assumption, however, and it will prove the weak link around which we will organize our criticism of the argument in the following section.

Our concern is that Cummins has overlooked important resources and strategies for calibration in general, and intuitions in particular. For not all calibration requires a certified basis that covers the whole target domain. Sometimes one can calibrate using a combination of information about some proper subset of the target domain on the one hand, and about the nature of the instrument and how it is meant to be able to track that domain on the other hand. We will call the first sort of information *the partial certified basis*, and the other *the theory of the instrument*, and together they license an *extrapolative inference* to the instrument’s accuracy over the target domain. We thus call this calibration strategy *extrapolative calibration*, which one might contrast with the *exhaustive calibration* strategy that Cummins considers. The partialness of the basis is essential to this strategy’s respecting the novelty requirement, and as we shall see, the theory of the instrument is essential to pulling off the inference at all.

Extrapolative calibration typically will start with a subset of cases from the target domain for which we know the correct value – the certified basis – and we test, diagnose, and correct the instrument on at least that portion. If the certified basis exhausts the target domain, then the novelty requirement will be failed, so it is important that we only have access to some sub-domain of the entire domain. Even given such a partial certified basis, we could still fail the novelty requirement if we were to perform a purely restrictive calibration at this stage, merely contracting the device’s target domain so that it becomes simply coextensive with the basis. In order to calibrate with novelty, we want to project out from that certified basis into territory that is not otherwise epistemically reachable. The extrapolative inference need not be an especially fancy one using high-powered statistics or the like. However, it is not frequently going to be as simple as the most flat-footed caricature of inductive inference, where one piles up a bunch of $F(a_i)$ and extrapolates out to $(x)Fx$. We should expect for any instrument – including intuition – that even at its best, it will have zones in which it will be error-prone, and bounds outside of which it cannot be expected even to be minimally reliable. So one can hardly hope for an unrestricted universal quantifier, and much of the game is figuring out what are the appropriate restrictions on it. A large part of doing so is learning what the discontinuities and nonlinearities are in the space of applications of the instrument. For example, using an iron spring scale to determine the mass of an object can be well-calibrated to operate near the earth’s surface in ordinary atmospheric and magnetic conditions, while

nonetheless being prone to errors at high altitudes or severe storm systems (due to reduced air pressure) or in the presence of a strong magnetic field (due to extra forces compressing or extending the spring). If you don't know anything about how the scale works, you might mistakenly over-infer from its successes in ideal conditions to its (chimerical) successes in sub-ideal conditions. Or, to put it differently: if you don't know anything about how the scale works, you won't be able to distinguish ideal from sub-ideal conditions in the first place.

Now, one does not necessarily need anything like a *complete* theory of how an instrument works, in order to calibrate it. One might discover that the scale gives odd results at high altitude while nonetheless lacking the physical theory that would predict and explain those oddities. But note that this may only be possible if one has some other, compensating resources. One might have a theory not so much of the device itself, but of the domain that the device is to measure: if we weigh a solid copper bearing at sea level, and are convinced that we have neither added to nor taken away any of its material as we climbed the mountain, then our theory of mass may tell us that its mass should be the same at altitude. Or one might have access to some other instrument for measuring mass, which one applies at the mountaintop to learn of the unwanted variation in results. So *some* degree of calibration is possible, even when one has only a very limited theory of the instrument. However, it will be very difficult to perform a rehabilitation on an instrument under such circumstances, for without knowing *why* it was that the device was giving screwy answers in the circumstances in question, it will be hard to know *what to do* to the device to prevent it from doing so in the future. We may have to settle, in such cases, for merely restricting our use of the instrument to the sorts of circumstances where the newly-revealed source of error, as revealed through testing, just cannot apply. So, ideally, one possesses both a fair amount of information about cases where the device has been made to give accurate results, and a sufficient theoretical account of how the devices's deliverances track different results of the target attribute in the world.

Cummins does not consider whether there might be sources of a basis other than a well-worked-out theory of the entire target domain. But we can think of three other possible sources for a partial certified basis. First, one might have an independently justified theory that only covers a proper subset of the target domain. Second, one might have other instruments or devices that yield results concerning some values within the domain. Third, there might a proper subset of the device's deliverances that one has independent reason to trust. We take the first two kinds of source to be self-explanatory; as for the last, note that our earlier story of the calibration of the telescope is of this type.

Cummins' line of reasoning could seem to apply to telescopes as far as towers are concerned, since we have non-telescopic access to their shapes and configurations. But when it comes to the moons of Jupiter or the rings of Saturn, no such independent information is available. Calibrating a telescope for astronomical purposes relies crucially on information about the workings of the device, in conjunction with getting it to give

expected results for towers and the like as well. It is a theory of the instrument that enables that crucial extrapolative step.

In summary, we find Cummins' argument overlooks the potential for extrapolative calibration, and underestimates the variety of resources potentially available for calibrating intuition. So philosophers ought not at this time preemptively foreclose on the possibility of an epistemically-fruitful calibration of intuition. Such a possibility, however, remains merely a *bare* possibility without some suggestion as to where a philosopher might acquire the relevant theory and or basis. In order to determine whether such a calibration may in fact be possible, we will have to consider what we currently know about the particular facts on the ground regarding intuitions; how they are produced, how they are used, and who has which ones.

Our next step is to think about what sort of resources are available for both aspects of an extrapolative calibration – the partial certified base and the theory of the instrument. While we think there is a lot of room for research on both these aspects it is our view that the second aspect (theory of the instrument) has been given shortest shrift. So in what follows we'll quickly cover some aspects of the state of play with regard to partial certified bases and then turn to think harder about a theory of the instrument.

4. Extrapolative calibration and intuitions

As we have seen, calibration may produce new epistemic value so long as the calibrator has both (i) sufficient sources for a partial certified basis (but not so generous as to fail to be partial), and (ii) a sufficient theory of the instrument to underwrite an extrapolative inference from partial certified basis to the projected performance of the device over its whole intended domain. Note that this is not a simple matter of having enough of two independent components, for the two factors must combine in the right way.

To evaluate the prospects for performing such an extrapolative calibration for intuition as a way of meeting the restrictionist challenge, we will first canvass the current prospects for contributions from non-intuitive sources, and then turn our attention to the question of intuitive resources for the calibration of intuitions.

4.1 Non-intuitional sources of philosophical evidence

In our initial presentation of the possibility of certification from a partial certified basis, we considered one way in which an independent theory can make possible domain-oriented calibration that satisfies the novelty requirement: namely, if the available theory covered only a proper subset of the target domain. One primary difficulty here, however,

is in finding a good candidate for such an independent-from-intuitions-but-also-partially-overlapping-with-them theory, and in a philosophical domain to boot. Scientific theories are an obvious place to look for such candidates¹⁵. If one could take a key category from a philosophical domain to be a natural kind, and then conduct a scientific investigation on it (rather than on people's reactions to it), one might begin to get such a theory.¹⁶

Another source for certifying a partial basis for calibration for intuitions would be other “instruments” than intuition itself, not so much in the sense of artifacts of measurement but rather in terms of other aspects of human psychology that might be thought to co-vary with some important philosophical facts. Intuitions and intuition-based beliefs, after all, are not the only part of human psychology that may be sensitive to the facts about moral goodness, knowledge, and so on. To the extent that various aspects of our biology may track matters of philosophical interest, we might be able to exploit that relation in order to gain a partial certified basis for the calibration of intuitions. Because such a relationship would not be consciously accessible, it would have to be determined experimentally. For example, assume (as is plausible) that humans are sensitive to the fairness of economic distributions in multiple ways – including, most importantly, in ways that do not always manifest in our intuitions, but which might be displayed behaviorally, perhaps even unconsciously. In concrete terms, we humans object if we regard our portion as unfair, a lack of complaint may therefore be regarded as evidence for the fairness of a distribution. In this situation, we may be able to design experiments to find out about fairness by simply operationalizing it in terms of the relevant behavior patterns. By experimentally instantiating various distributions and seeing how subjects react, we could get non-theoretical access to the target domain without relying on reports of intuitions about fairness. Perhaps research using the ‘ultimatum game’ would be useful start here; see, e.g., Skyrms (1996); Nowak et al. (2000).

It is unlikely that such experimental access will be possible for all cases where issues of fairness arise. We may take it therefore, at least provisionally, that this sort of access is limited, thereby providing a partial certified basis from which the extrapolative inference could be made (assuming, of course, that PI's can be successfully calibrated against that partial basis in the first place). But that's just one possible link between experiment and intuition. Our affective and normative systems are tightly wound around each other; so eye dilation, skin conductance, cortisol levels, and even detectable patterns of neural activity may reveal moral responses which may then offer further data for calibrating relevant intuitions; see, e.g., Cima *et al.* (2010) on moral judgment in psychopaths. Brain imaging technologies may also prove relevant in this regard (e.g., as reviewed in Greene & Haidt (2002)).

Where such theories or scientific results can be found, we should look to exploit them. Nonetheless we fear that their contribution, albeit real, will remain minimal. For most areas of philosophy, one simply does not, and probably will never, find empirical theories

that overlap the target domain. There does not look to be an empirical science of, say, metaphysical necessity, or the moral good. And there is a further danger here that the theory in question may so radically disagree with our intuitions about the matter that it yields not a certified basis for, but rather a thorough *decertification* of the relevant intuitions. The lesson from modern physics is not, after all, that our intuitions about the structure of space and time are reliable, nor even how we might go about making them better, but basically that we're better off just leaving our intuitions behind altogether.

4.2 *Calibrating intuitions on other intuitions?*

To the extent that we are going to calibrate intuitions in any sort of wholesale fashion, we are going to need to do so in large part on the basis of our best philosophical instrument, which unfortunately is intuition itself. This, needless to say, makes things complicated. For we can use intuition as a successful calibrator for intuition only provided we have some reason to extend a heightened degree of trust to some *subset* of intuitions under consideration. The list of candidate subsets already on offer is considerable: ordinary intuitions, consensus intuitions, reflective intuitions, expert intuitions and clear/forceful intuitions just to name some of the most obvious contenders. What would provide us with such a reason would be some version of a theory of the instrument for intuition – an account of some sort as to how intuition works, that can inform decisions as to where it may have greater or lesser susceptibilities to the sorts of errors that restrictionists have made much of.

At first glance it may seem clear that any theory of the instrument for intuition is going to commit us to doing a great deal of hard core cognitive psychology, and since the relevant work is not yet done, we are in no position to theorize about the nature of intuition.¹⁷ But in fact, that's not the only way to generate a workable theory of the instrument for intuition. For at least some kinds of instruments one can know a fair amount about their effectiveness under various sorts of circumstances, without having first to possess a detailed knowledge of their mechanism. For example, folks who grow up in cold climates know we need about twelve times more distance to brake our cars in the snow and ice than we need under ordinary road conditions. Furthermore they know this without having, or at least without needing to have, any real handle on the mechanism(s) that brake our cars. It behooves us then to think about the possibility that philosophers and their use of intuition may be like drivers in temperate climates and their use of brakes: we can know a great deal about it – in particular, how to use it in various circumstances – without grasping all that much about the particulars of its mechanism. In short, do we have a *folk* theory of the instrument for intuition, appropriate for grounding at least some extrapolative calibration of intuition?

Our view is that, as things stand, all the major contenders mentioned above suffer from one of two shortcomings. Some proposals do seem to provide the material for a genuinely trustworthy partial basis, but that basis turns out to be insufficiently ambitious to do any real philosophical work. We argue below that intuitions about ordinary or easy cases, or intuitions that meet with a widespread consensus, will all be have this difficulty. We will thereupon contend that those approaches that purport to have sufficient strength to do enable substantive philosophical progress unfortunately do not seem, as a matter of fact, to provide any certification. That is, we have no reason to think that the subset they mark out do not stand in need of the very calibration they are supposed to help provide.

4.3. The Ordinary, The Easy, & The Consensus: You Can't Get There From Here – Although Here is a Lovely Place to Be!

Perhaps the most worked out example of an appeal to widely accepted intuitions is to be found in the work of George Bealer (see for example Bealer 1996). On Bealer's account, although intuitions may be fallible (he admits that strongly felt intuitions (such as the naive comprehension axiom) may be erroneous), he nonetheless thinks that we have evidence that they are “for-the-most-part” correct (Bealer (1996)). He bases his trust of intuitions on the “on-balance consistency of our elementary concrete-case intuitions” (Bealer (2000)). Let us grant that evaluation at least for the sake of argument, though it strikes us as highly plausible in its own terms. The problem, in terms of calibration, is that we are in a position of fairly substantial ignorance at this time as to whether intuitions *about the sorts of cases philosophers often consider* is of a sort that shares in the same sources of on-balance high reliability that can be found with our most ordinary and clear cases. The folk theory of the instrument for intuition that Bealer appeals to does not give us any means to extend our trust in ordinary intuitions to the intuitions of use to philosophers.

In fact, there are several good reasons to worry that this is not so – that in fact, performance on quotidian cases might not be a good predictor of performance on cases more typical of philosophical discourse. First, recent work in psychology raises worries about the very idea that our performance in one intuitive area can shed light on our performance in others. For many psychologists and philosophers¹⁸ have come to believe that we do not have one big domain-general intuition system, but rather a number of distinct domain-specific mechanisms subserving our cognition in these areas—what Steven Pinker has colorfully called “the mind as a Swiss Army knife” (Pinker (1997)). If such ‘massive modularity’ theories are correct, then our verifiable successes in one domain (like the everyday world of middle-sized dry goods) would, at best, be evidence that the intuition-producing mechanism subserving *that* domain can be trusted. But just as evidence that your thermometer is reliable is no evidence that your barometer is working, so too would evidence that the mechanism for one domain is reliable fail to be evidence

for the reliability of a different intuition-producing mechanism operating in a different domain. This is not to say that we are committed to any form of massive modularity turning out true, and we recognize that it is still a matter of significant debate in cognitive science. Rather, our concern is that in the absence of compelling reasons to think that all intuitions – both ordinary and esoteric – arise from a unified psychological source, no easy extrapolative step will be justifiable.

A second worry is the suspicion that the sorts of cases philosophers are wont to deploy in their arguments will have a tendency to be of a sort where ordinary sources of reliability would be expected to break down. With ordinary cases of determining that whether *A is F*, we might expect that the various factors that are of primary relevance to determining the categorization of things as *F* or not-*F* will largely be in agreement regarding *A*. Typical non-ducks will neither look like a duck, nor swim like a duck, nor quack like a duck. Yet philosophers often need to recruit cases where this consilience exactly breaks down. Epistemologists want to know what happens when availability of internal reasons and reliability of external causes of belief come apart. Ethicists want to know what happens when actions that will likely produce the greatest good for the greatest number are distinct from the actions that fully respect the autonomy of human persons. And so on. We often want to consider cases that look just like ducks, quack just like ducks, but swim just like otters on steroids; it would be unsurprising if our ordinary reliability in duck/non-duck judgments break down in such cases.¹⁹

Furthermore, there is good reason to worry that the problematic viability of extrapolation from easy cases to hard ones is not just a bare possibility. For example, our intuitions about the comparative cardinality of finite sets – e.g., that taking every other member from a set yields a set half as big – seem unproblematic, and we do not doubt they are generally well-calibrated. But we do doubt that this calibration extends into the area of infinitary mathematics. We know that many principles that hold in finite domains do not hold in infinite ones, and so we may legitimately object that our confidence for judgments about *n*'s cannot be extended to our judgments about \aleph 's. This, of course, does not stop us from making use of our intuitions in inappropriate ways. For example, BonJour appeals to the intuition that there are half as many even integers as there are integers (BonJour 1998, 209n24). Similar situations arise if we look to move from arithmetic to trickier areas of mathematics. Whatever intuitive success we may legitimately have with basic addition and multiplication cannot justify any expectation that we would meet with similar success in probability theory, where in fact people generally have appallingly bad intuitions. So it does not seem that any general licensing of the difficult by the easy will be forthcoming.

Very similar sorts of concerns apply to trying to use intuitions for which there is consensus as a partial certified basis. Recall that calibration involves making the device's reports fit with what is known through the certified basis. Such a comparison can only be

done in cases covered by both the device's reports and the certified basis. But with consensus, the certified basis consists of cases where the reports of everyone's device agree. This means that *any* relevant device's reports – in this case, any competitor set of intuitions under particular circumstances – will trivially fit with the certified basis, since that is part of what it is to be part of the certified base in this case (i.e. our certified basis is one that is, *prima facie*, equally easily generated by devices that stand, in other circumstances, in substantial disagreement). So we need some grounds for giving preference to one projection from the consensus data to another one. Perhaps this could come from some of the other sources considered above, at least to some extent. But it is also just the sort of situation where a theory of the instrument could do a lot of work. But that is just what is lacking on this approach. As Cummins argues:

We might try to use only uncontroversial cases – that is, cases everyone agrees on in advance. But if we know that everyone, including our subject [i.e., the intuitor being calibrated], agrees on the test items, there is no point in administering the test. If the subject gives a “wrong” answer, that just shows that the item did not belong on the test. Of course, we could simply be looking to see if the subject is like everyone else we have tested so far. This might be interesting. Perhaps, pursuing this strategy, we might find that everyone, or nearly everyone, or nearly everyone in a certain culture, or economic class, or what have you, shares certain fairness judgments. That would be worth knowing, but it would not count as calibrating our subject's fairness intuition. ((1998), p. 117)

For example, suppose that a Kantian and a utilitarian were attempting to calibrate their intuitions against consensus. For both, their agreed-upon intuitions are thus trivially certified. But we would need some way to extrapolate out from the consensus cases in order to settle which is correct where they disagree. A certified basis made of easy-case intuitions cannot, without substantial further assistance, allow the Kantian or the utilitarian to extrapolate the accuracy of their respective hard-case intuitions, since each has an equal claim to comporting well with the basis. Consensus cases, by their very nature, will probably fail to select one side over the other²⁰. We think that Cummins point is on target, but does not go as far as he thinks it does. The real lesson here is not that consensus cases cannot be of use in calibration. It is that they cannot be of use in calibration *without the assistance of a theory of the instrument that will license particular extrapolations*. A theory of the instrument that does *not* arise naturally from a focus on consensus intuitions and that does not appear, at least to this point, to have been generated independently by those promoting the value of consensus.

So much for subsets of intuitions whose members are plausibly taken to be trustworthy, their problem is the absence of a ‘theory of the instrument’ even a folk one, that will

ground the extrapolation of the trustworthiness of their basis to a majority of the intuitions at work in philosophical contexts.

4.4 *The Vagaries of Clarity, Reflection, & Expertise*

The other members of our list of candidate subsets of intuition: clear/forceful intuitions, reflective intuitions, expert intuitions, suffer from a different problem. Their candidate theories of the instrument do generalize to philosophical contexts; the problem is with the claim of the relevant subsets to genuinely trustworthy.

The folk theory of intuition with the longest history is probably the view that genuine or trustworthy intuitions have a phenomenal quality that marks them off from all other intellectual seemings. While it's not clear that this is a view that gets much play in the current literature, it has a distinguished history. The Stoics' criterion of truth perhaps fits this picture as, of course, does the thought of many Early Modern thinkers, most famously Descartes. Obviously there is more to be said, but for now we'll settle for pointing out that this approach has always struggled with two problems. First, what seems clear and distinct varies from thinker to thinker. Second, that which is accepted as clear and distinct by one age has not always fare well in the longer run. In our view this pair of problems rules out the 'clear and distinct' approach as a workable theory of the instrument for intuitions.²¹

More popular in recent years is the idea that intuitions that are the product of greater *reflection* can be expected to be trustworthy, and indeed will be sheltered from the kinds of effects seemingly uncovered in the restrictionists' experiments. Philosophers proposing this idea typically make little or no appeal to any systematic evidence that reflection has this inoculating capacity; it is the folk theory that they are relying on. Such an approach will obviously apply to the intuitions at work in philosophical contexts, but are such intuitions really trustworthy? We will focus on a version of this found in Laurence Bonjour, which we take to be typical in its features.²²

While admitting the fallibility of rational insight, Bonjour offers a method for identifying those intuitions which are faulty: careful reflection. Bonjour distinguishes between mistakes that are externally and internally correctable, where the latter require only thinking harder about or re-examining the problem. Bonjour claims it is possible to identify faulty intuitions by "appeal[ing] to the fact that many such errors, and perhaps all of them, are correctable 'from the inside' via *further reflection*" (Bonjour 1998 p. 117, emphasis added).

If it is true that further reflection is all that is ever needed to identify which intuitions need to be rejected—if their falseness becomes manifest through simple re-examination,

then BonJour has provided analytic philosophy with all it may need by way of addressing the troublesome evidence. But BonJour's primary evidence for his claims are themselves based on intuition. We grant BonJour that many cases of divergent intuitions can be rectified by closer inspection, such as "most routine mistakes of reasoning or calculation, which yield to equally routine corrections" (*ibid.*, 116-117). But as he notes, "[t]he important question is whether *all* mistakes of apparent rational insight, all cases in which something seems necessary that is not really necessary, are mistakes that are internally correctable in this sense. Such a thesis *seems plausible enough on an intuitive basis*, though there is no apparent way to argue for it..." (*ibid.*, 117, latter emphasis added).

BonJour's defense of the claim that further reflection reveals which intuitions to discard is fraught with problems. The reason we can take ourselves to be able to identify faulty intuitions through reflection, according to BonJour, is that we have the intuition that we can. But we would contend that *that* intuition is itself contested, and not one that can be relied upon at this stage. It is obviously not an intuition endorsed by opponents of intuition, and indeed the authors of this paper feel compelled to report that we fail to share it, too.²³

Moreover, we disagree with BonJour's assertion that this claim is one for which only intuition could serve as evidence. Numerous empirical tests are available. One obvious sort of test is induction from past successes and failure. But, despite a number of local successes, the persistent divergence of intuitions across philosophers (all of whom are surely 'thinking hard' on their thought-experiments!) would seem to indicate that this is not a method with a sufficiently high rate of success in weeding out incorrect intuitions. One could also conduct controlled experiments in which subjects were given varying incentives; or different instructions; or that tested the subjects' intrinsic motivation to think hard, known in the psychology literature as 'need for cognition'.²⁴ Just how easily and reliably one can get subjects to perform at ceiling on which sorts of intuitive judgment tasks – of more or less esoteric sorts as well – is an empirical question, and one which should not be begged with a bald assertion that it falls under intuition's purview. Surely *some* tasks can be facilitated by just thinking about them more clearly.²⁵ But there is also evidence that thinking too hard can also cloud one's judgment and 'lock in' an early mistake – especially in conditions in which the relevant information is neatly packaged, as is often the case with philosophical thought-experiments (i.e., one is supposed to think only about the contents of the scenario, and only with regard to the question of whether it is or is not a case of knowledge, free will, or the like).²⁶ Which kind of influence would be the dominant influence on intuitions subject to reflection is simply an empirical question that the philosophical community currently lacks any information on, and thus we cannot at this time assume that the "think harder" approach will bear much calibrational fruit. It likely will bear *some*, but at this time we do not know quite where to find it. Only once it ceases to be a folk theory of the instrument of

intuition, and becomes part of a scientific one, can the appeal to reflection do philosophers much good.

The last folk theory of the instrument we will consider here is that of expertise. It might be suggested that we could find out which propositions the experts in a given area of philosophy endorse, and let those form the basis for calibration. The line that philosophers are on the whole experts is, as we noted briefly in our introduction, one popular approach for attempting blocking the restrictionist challenge in the first place – for if philosophers are already relevantly expert, then no calibration of their intuitions would be needed. Our comments here are not directed at those arguments, but rather at a different line of thought: perhaps some elite *subset* of philosophers can be certified trustworthy in a way that the rest of us cannot, but off of whose intuitions the rest of us can attempt to calibrate – a kind of Greenwich Mean Philosopher. Experts presumably have access to the truths of their area of expertise, so bringing our intuitions in line with experts' endorsements looks like an attractive method of calibration. An examination of the possible evidential sources for experts' judgments, however, reveals problems with using them as the basis for calibration. Someone endorsing this move faces a dilemma: experts base their judgments either on their own intuitions, or on something else. If experts base their judgments on their philosophical intuitions about the target domain, then we may legitimately ask how *those* intuitions were themselves calibrated. And the answer may, again, be based either on intuitions, or on something else. Given the problem at hand – how to calibrate intuitions, given that the practice of appeal to intuitions is now in a state of challenge – this cycle cannot stop at simply more intuitions. For in that case, our answer to the problem of how we might calibrate would presuppose that the problem was already solved.

So at some point in this iteration, experts must base their judgments on something other than intuitions, such as independent theories or experimental data. But if this were so, then there would be no need to appeal to expert intuitions at all, since we might just as well calibrate intuitions off of whatever it is that the experts were basing their judgments on. To play on Wittgenstein, the appeal to expert intuitions in calibration would be at best an unnecessary shuffle.²⁷

We thus set aside the idea that appeals to experts can, at this time, be of any use in pursuing a project of calibration.²⁸

4.5 Is one philosopher's poison another philosopher's meat?: The case of "theory contamination"

We close this section with a case study in the difficulty of putting forward a useable folk theory of the instrument, even when it is meant only to do fairly limited and narrowly

defined work. Philosophers have been aware of at least some of the methodological dangers of intuitions, and one can find moments in the literature in which they claim to identify intuition-corrupting influences at work in certain circumstances, and based on this identification reject the intuitions generated under such circumstances. After the corrupted intuitions are rejected we are meant to be left with a set of intuitions bereft of conflict, properly attuned to the target domain, and therefore apt for deployment as the evidential basis for the relevant philosophical theory. Just such an approach is set out explicitly in Goldman and Pust (1998), when they consider the possibility of “ill-informed or misinformed” subjects (183), or “theory contamination,” which threatens when “the person experiencing the intuition is a philosophical analyst who holds an explicit theory about the nature of F, [and as a result] this theory might warp her intuitions about specific cases” (ibid.). The theory-corrupted intuitions are thus rejected, and the pre-theoretical ones preserved.

We suggest that this kind of strategy is best seen as an example of extrapolative calibration, taking whatever mechanism or mechanisms that produce PI's to be the instrument in question. But this strategy is one that relies more centrally on a theory of the instrument than it does on having an extensive partial certified basis to build on. Goldman and Pust's philosopher attempts to identify those factors that might influence the proper working of the intuition-producing mechanisms by means of an appeal to some *theoretical* account of the relevant sources of intuition corruption. The problem for the philosopher is that he must do this with little knowledge of how those mechanisms work. The philosopher faces the challenge of attempting extrapolative calibration without actually knowing very much about the instrument that she hopes to calibrate. As Timothy Williamson has written, “‘Intuition’ plays a major role in contemporary analytic philosophy's self-understanding. Yet there is no agreed or even popular account of how intuition works, no accepted explanation of the hoped-for correlation between our having an intuition that P and its being true that P” (Williamson (2008), p. 215). This may overstate matters somewhat – but only somewhat. Considering the Goldman and Pust strategy, for example, one faces the question of why it is that those intuitions that are influenced by theory are thus *contaminated*, since one might just as well claim that these intuitions are the ones that are *well-informed*, and are thus *more* likely to be trustworthy than the untutored inclinations of the *hoi polloi*. (See McBain (1999) for a forceful defense of this latter possibility.) This kind of calibration needs not just a proposed factor to separate wheat from chaff, but moreover some capacity for allowing us to know which pile of vegetable matter to discard at the end, and from which to bake our good philosophical bread.

Perhaps this problem will be made clearer through a familiar example. Going back at least to Plato and Kant, and continuing more recently through such theorists as Kohlberg (1969), many ethicists have valued judgments untainted by emotion, for fear that irrational emotion may corrupt our intuitive rational apprehension of the moral good. As

uncontroversial as this sounds, it actually involves unfounded assumptions about the nature of the intuition-producing mechanisms. Given that we don't know where intuitions come from, what reason do we have for thinking emotional attachment to a case would adversely affect one's intuitive mechanism? It could just as easily be the case that affect *positively* influences the accuracy of one's intuitions.²⁹ The problem is that without a rich, constructive theory of the instrument for philosophical intuitions, we have little means for discerning what particular factors are corrupting those intuitions. Without an account of the device's proper functioning, it simply isn't possible to identify which cases are the ones where something is interfering with that proper functioning.³⁰

Overall, then, our quick survey of folk theories of the instrument for intuition suggest that intuitions are not as much like cars as we might have hoped. Absent a good account of the mechanism we are in dark with regard to how to extrapolate from regions of successful intuition, though we do think that one can legitimately point out limited regions where success can be expected. This suggests that, minimally, we need to come up with better folk theories of intuition. Failing that we will have to fall back on the full dress cognitive psychology approach. This in turn suggests both how much work remains to be done before philosophers can really rely on intuitions but also that there is reason to think that such work could be done.

5. Conclusion

Our investigations leave philosophers in a something of a bad-news, good-news situation. The first bad news is that calibration cannot at present help philosophers respond to the growing challenges to intuition-based methods. The good news is that calibration *may* be of some use to the analytic philosopher in the future, if only philosophers can successfully pursue one or more of the avenues listed above. We meant to open up an avenue of hope in section 3, in our response to Cummins, that we did not mean to shut down in section 4 – rather, we meant to show only that it is an avenue that has yet been inadequately explored.

However, there is a further bad-news punch line here, of a metaphilosophical sort: in order to take this route of saving the practice of using intuitions, we must either turn to science for serious help in exploring the extent of intuitive consensus and, moreover, the structure of our intuition-producing mechanisms; or develop other alternative methods within philosophy that do not themselves rely on intuition. Calibration offers philosophy no way to save its intuitions without looking to use those intuitions, to a significant degree, in a way that may completely forsake any claims of methodological autonomy for them.

Bibliography

- Alexander, J. and Weinberg, J. (2007). "Analytic Epistemology and Experimental Philosophy." *Philosophy Compass*, 2(1), 56–80.
- Alston, William. (1993). "Epistemic Desiderata", *Philosophy and Phenomenological Research* LIII (3): 527- 551.
- Bealer, George. (1996). "On the possibility of philosophical knowledge," *Philosophical Perspectives* 10: 1-34.
- (1998). "Intuition and the Autonomy of Philosophy". In DePaul and Ramsey.
- (2000). "A theory of the a priori," *Pacific Philosophical Quarterly* 81:1-30.
- Bechtel, W. (2000). "From imaging to believing: Epistemic issue in generating biological data." In R. Creath and J. Maienschein (Eds.), *Epistemology and Biology*, 138-163. Cambridge: Cambridge University Press.
- Bishop, M. and Trout, J. D. (2005). *Epistemology and the Psychology of Human Judgment*. Oxford: Oxford University Press.
- Blackburn, S. (1998) *Ruling Passions*. Oxford: Oxford University Press.
- Boghossian, P. and Peacocke, C., eds. (2001). *New Essays on the A Priori*. Oxford: Clarendon Press.
- BonJour, Laurence (1985) *The Structure of Empirical Knowledge*. Cambridge, MA: Harvard University Press.
- (1998), *In Defense of Pure Reason*. Cambridge University Press: Cambridge U.K.
- Castelli, F., Happé, F., Frith, U. and Frith, C. 2000. "Movement and mind: A functional imaging study of perception and interpretation of complex intentional movement patterns." *Neuroimage*, 12, 314 – 25.
- Chang, Hasok (2004), *Inventing Temperature: Measurement and Scientific Progress*. Oxford; Oxford University Press.
- Craig, Edward (1990). *Knowledge and the State of Nature*. Oxford: Clarendon Press.
- Cummins, Robert (1998) "Reflection on Reflective Equilibrium," in DePaul and Ramsey, 113 – 128.

- DePaul, M. and Ramsey, W. (eds.) (1998) *Rethinking Intuition: The Psychology of Intuition and its Role in Philosophical Inquiry*. Lanham, MD: Rowman and Littlefield.
- Descartes, R. (1984) *Meditations on First Philosophy* in John Cottingham, Robert Stoothoff, Dugald Murdoch, (translators) *The Philosophical Writings of Descartes*. Cambridge: Cambridge University Press.
- Dretske, Fred (1970), “Epistemic Operators”, *Journal of Philosophy*, 67: 1007-23.
- Evans, J. and Over, D. 1996. *Rationality and reasoning*. Hove: Psychology Press.
- Feltz, A. (2008). “Problems with the appeal to intuition in epistemology.” *Philosophical Explorations*, 11, 131 – 141.
- Fix, John D. (1995) *Astronomy: Journey to the Cosmic Frontier*. St. Louis: Mosby.
- Franklin, A (1997). “Calibration.” *Perspectives on science* (1063-6145), 5 (1), p. 31.
- Gettier, E. 1963. “Is Justified True Belief Knowledge?” *Analysis* 23: 121-123.
- Gibbard, A. (1990) *Wise Choices, Apt Feelings: A Theory of Normative Judgment* Oxford: Oxford University Press.
- Goldman, A. and Pust, J. 1998. “Philosophical Theory and Intuitional Evidence.” In Michael R. DePaul and William Ramsey (eds.) *Rethinking Intuition: The Psychology of Intuition and its Role in Philosophical Inquiry*. Lanham, MD: Rowman and Littlefield, pp. 179- 200.
- Greene, J. and Haidt, J. 2002. “How (and where) does moral judgment work?” *Trends in Cognitive Sciences*, 6(12), 517-523.
- Guski, R. and Troje, N. 2003. “Audiovisual phenomenal causality.” *Perception & Psychophysics*, 65(5), 789 – 800.
- Hirschfeld, Lawrence A. and Susan A. Gelman (eds.) (1994) *Mapping the mind: Domain specificity in cognition and culture*. New York, NY: Cambridge University Press.
- Jackman, H. (forthcoming). “Intuitions and semantic theory.” *Metaphilosophy*.
- Kahneman, D., P. Slovic, and A. Tversky (eds.) (1982) *Judgment under uncertainty: Heuristics and biases*. New York, NY: Cambridge University Press.
- Katz, J. 1998. *Realistic Rationalism*.

- Kauppinen, A. 2007. "The Rise and Fall of Experimental Philosophy", *Philosophical Explorations* 10(2): 95-118.
- Kohlberg, L. (1969). "Stage and sequence: The cognitive-developmental approach to socialization." In D. A. Goslin (Ed.), *Handbook of socialization theory and research*, 347-480. Chicago: Rand McNally.
- Knobe, J. (2003). "Intentional action and side effects in ordinary language." *Analysis* 63, 190 – 193.
- (2004). "Intention, intentional action, and moral considerations." *Analysis* 64, 181 – 187.
- and Burra, A. (forthcoming). "Intention and Intentional Action: A Cross-Cultural Study." *Journal of Culture and Cognition*.
- Kornblith, Hilary (2002) *Knowledge and its Place in Nature*. Oxford University Press: Oxford U.K.
- Lackey, J. (forthcoming) "Learning from words." In *Philosophy and Phenomenological Research*.
- Lehrer, K. (1971) "How reasons give us knowledge, or, the case of the gypsy lawyer." *Journal of Philosophy*, 68(10), 311-3.
- and Cohen, S. (1983). "Justification, truth, and coherence." *Synthese* 55, 191 – 207.
- Levin, J. (2004). "The evidential status of philosophical intuitions." *Philosophical Studies*, 121(3), 193-224.
- Loftus, E. (2003). "Our changeable memories: legal and practical implications." *Nature Reviews: Neuroscience*. 4(3), 231-4.
- Machery, E., Mallon, R., Nichols, S., and Stich, S. P., (2004), "Semantics, Cross-cultural Style," *Cognition*, 92(3): B1-B12.
- McBain, James. (1999). "The Role of Theory Contamination in Intuitions," *Southwest Philosophy Review*, 15(1):197-204.
- Miller, R. (2000). "Without Intuitions". *Metaphilosophy*, 31, 231 – 50.
- Nahmias, E., Morris, S., Nadelhoffer, T, and Turner, J. (forthcoming). "Surveying Freedom: Folk Intuitions about Free Will and Responsibility". *Philosophical Psychology*.

- Neta, R. (forthcoming) "Epistemology Factualized: New Contractarian Foundations for Epistemology". *Synthese*.
- Nichols, S. (2004). *Sentimental Rules*. Oxford: Oxford University Press.
- and Knobe, J. (forthcoming) "Moral responsibility and determinism: The cognitive science of folk intuitions." In *Nous*.
- , Stich, S., and Weinberg, J. (2003). "Metaskepticism: Meditations in Ethno-Epistemology." In *The Skeptics*, ed. S. Luper. Burlington, VT: Ashgate.
- Nisbett, Richard and Ross, Lee (1980). *Human Inference: Strategies and Shortcomings of Social Judgment*. Prentice-Hall.
- Nowak, M., Page, K., and Sigmund, K. 2000. Fairness versus reason in the ultimatum game. *Science*, 289(5485), 1773-1775.
- Peacocke, C. (2001). Explaining the a priori: The programme of moderate rationalism. In Boghossian & Peacocke (2001), 255 – 85.
- Pinker, S. (1997) *How the Mind Works*. Penguin Press.
- Pust, J. (2000). *Intuitions as evidence*. New York and London: Garland Publishing
- Samuels, R. (1998). Evolutionary psychology and the massive modularity hypothesis. *British Journal of Philosophy of Science*, 49: 575-602.
- (2000). Massively modular minds: evolutionary psychology and cognitive architecture. In P. Carruthers and A. Chamberlain (eds.), *Evolution and the Human Mind*. Cambridge University Press.
- Sosa, E. (1996) Rational intuition: Bealer on its nature and epistemic status. *Philosophical Studies*, 81(2-3), 151-62.
- (1998). Minimal intuition. In Depaul & Ramsey (1998).
- Sperber, D. (2002). In Defense Of Massive Modularity. In Dupoux, E. *Language, Brain and Cognitive Development: Essays in Honor of Jacques Mehler*. Cambridge, Mass. MIT Press. 47-57.
- Skyrms, Brian (1996). *Evolution of the Social Contract*. Cambridge: Cambridge University Press.

Stanovich, K. (1999). *Who is Rational?: Studies of Individual Differences in Reasoning*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Stich, S. (1988). “Reflective equilibrium, analytic epistemology and the problem of cognitive diversity”, *Synthese* 74, 391-413.

Strong, Frederick. (1999) “Guest editorial - The History of the Double Blind Test and the Placebo”, *J. Pharm. Pharmacol.*, 51, 237-8.

Swain, S., Alexander, J. and Weinberg, J. 2006. The instability of philosophical intuitions: Running hot and cold on Truetemp. Presentation to the annual meeting of the Society for Philosophy and Psychology.

Tidman, Paul (1996), “The Justification of the A Priori Intuitions”, *Philosophy and Phenomenological Research* 56:1 pp. 161 – 171.

Weatherson, B. (2003) What good are counterexamples? *Philosophical Studies*, 115: 1-31.

Weinberg, J. (forthcoming) What’s Epistemology For?: The Case for Neopragmatism in Normative Metaepistemology. In S. Hetherington (ed.), *Epistemology Futures*. Oxford: Oxford University Press.

-----, Nichols, S., & Stich, S. (2001). Normativity and epistemic intuitions. *Philosophical Topics*, 29(1&2), 429–459.

Williamson, T. (2004a). “Philosophical ‘Intuitions’ and Skepticism about Judgement”. *Dialectica*, 58(1), 109 – 53.

----- (2004b) “The Presidential Address I—Armchair Philosophy, Metaphysical Modality and Counterfactual Thinking. *Proceedings of the Aristotelian Society* 105 (1), 1-23.

¹ E.g., Weinberg *et al.* (2001) , Swain *et al.* (2008), Machery *et al.* (2004), Nichols & Ulatowski, Zamzow & Nichols (2009)); Feltz (2008); Feltz & Cokely (2009); Liao *et al.* (forthcoming); Machery (2011); Weigel (forthcoming).

² For recent examples, see Weinberg *et al.* (2010).

³ See Weatherson (2003); see also Crowley & Weinberg (2010) for a response.

⁴ This phenomenon is called ‘chromatic aberration’; see King p. 44.

⁵ See King, chapters 8 and 14.

⁶ In addition to the approaches to chromatic aberration discussed here there was also the whole move to reflecting telescopes along the lines suggested by Newton (see King for details).

⁷ Not too much has been published in the philosophical literature on calibration, but an important exception is Franklin (1997).

⁸ We will henceforth use “device” instead of “putative source of evidence,” trusting that nothing here is lost by doing so. We recognize that not all sources of evidence can be usefully assimilated to devices – to do so with testimony would be to lose our grasp on the epistemic agency of our fellows, for example. But there is no such loss at risk with intuitions. Moreover, we take it that calibration will simply be a non-starter for someone who takes there to be a radical gulf between intuitive judgments on the one hand and scientific instruments on the other. So this paper is aimed more at those willing to endorse a fundamental similarity between the two.

⁹ “Accuracy over a domain” here should be understood somewhat broadly, so that it also includes degrees of precision. One thermometer may be better than another because it can accurately record higher or lower temperatures; but also one thermometer may be better than another because it can accurately make more fine-grained temperature discriminations. We take it that all the points made throughout this discussion that are in terms of areas or domains will apply, *mutatis mutandis*, to degrees of precision.

¹⁰ For a rich and thoughtful account of the full range of challenges that must be overcome in order to calibrate a scientific instrument, with a particular focus on thermometers, see Chang (2004).

¹¹ Although this is the same general idea of “restriction” that gives the restrictionist challenge its name, we would emphasize that one can be in favor of the idea of restrictionist calibration of intuitions *without* thereby being committed to the restrictionist arguments more generally. For example, ordinary language philosophers might take themselves to have good reasons to want to restrict philosophers’ intuition-deploying practices only to very quotidian cases. But they might nonetheless think that this restriction is what that can be both motivated and carried out without leaving the armchair.

¹² Of course, in correcting an information-gathering device, some combination of restriction and rehabilitation might be called for; we might call this sending the device for some much-needed ‘R&R’.

¹³ For more background on the phenomenon of seeing and astronomers’ attempts to avoid it, see Fix (1995).

¹⁴ There is no need to deny, and Cummins does not deny, that one can still calibrate a device when the novelty requirement is not fulfilled, for various non-epistemic reasons. It may be useful, for example, if the calibrated device will ultimately be more convenient to use than other available means. For some sufficiently large look up table and some sufficiently small calculator there is complete overlap in domain, but it is clear that the calculator is pragmatically superior. In such instances, however, calibration will not be epistemically worthwhile: it will not extend our knowledge.

¹⁵ We should note that philosophers do not have to turn to scientists in order to find fairly intuition-free theories of their target domains. Edward Craig’s recent account of knowledge in his (1990) does not depend centrally on esoteric intuitions as evidence, in that it appeals more to the role that a concept like knowledge

can play in our overall cognitive economy. However to this point Craig's work appears to be something of a methodological orphan.

¹⁶ A paradigm example of this approach is Hilary Kornblith's (2002) work on knowledge. While we do not necessarily endorse the details of Kornblith's account of knowledge, we encourage investigation into the viability of researching objects of philosophical inquiry as natural kinds. If such approaches work, they could provide a welcome further source for a certified basis.

¹⁷ Talbot (2010) argues for the importance of philosophers learning the underlying psychology of intuition, as an aid to philosophical methodology.

¹⁸ E.g., Samuels (1998, 2000); various papers in Hirschfeld and Gelman (1994); Sperber (2002); various papers in Carruthers and Chamberlain (2000); Carruthers (2004, 2006); Barrett (2005). Note that the relevant notion of modularity here is only that of functional specialization, and need not include the full suite of Fodorian modular characteristics like automaticity or cognitive impenetrability (see Barrett and Kurzban (forthcoming)).

¹⁹ The results in Wright (2010) can be seen as supporting this contention.

²⁰ A similar point is made regarding the analysis of justification by William Alston in his paper 'Epistemic Desiderata' p. 537 in particular (Alston 1993).

²¹ This phenomenological approach has had its more recent proponents, most notably George Bealer. See Weinberg & Alexander (forthcoming) for a response.

²² See also Kauppinen (2007); Sosa (2007).

²³ In fact, at least some of us find the opposite intuitive: that we can't generally identify faulty intuitions via intuition itself.

²⁴ Cacioppo and Petty (1982); Cacioppo, Petty, and Kao (1984)

²⁵ Stanovich and West (2000); Epley and Gilovich (2005)

²⁶ Petty *et al.* (2001). See Pinillos *et al* (2011) and Weinberg *et al* (forthcoming) for some specifically philosophical results along these lines, i.e., that sometimes reflection can help, and other times, not so much.

²⁷ We would also note that this entire discussion is setting aside the fundamental question of deciding who counts as a relevant philosophical expert. We can well imagine those who have studied closely the texts of Descartes or Russell claiming an expert status on the topic of knowledge, for example, that students of Peirce or Austin might wish to deny. Who gets to count as an Officially Licensed Intuitor? If the epistemic status of philosophical intuitions turns out to be a matter of who can win the relevant political battles over credentialing, then the postmodernists will truly have won.

²⁸ There is a certain structural similarity between this argument and Cummins' main argument as discussed above, so the reader may wonder whether this argument might be vulnerable to the same kinds of objections we lodged against Cummins. However, a key difference is that Cummins is there aiming to

show that intuitions can never be calibrated except when it would not be useful to do so; whereas here we are trying to make the much more limited and targeted argument that the restrictionist challenge cannot be met by means of calibration with appeal to experts' intuitions. We should note nonetheless that this section of owes a substantial debt to his discussion.

²⁹ See, e.g., Gibbard (1990); Blackburn (1998); Nichols (2004).

³⁰ Weinberg, Nichols and Stich (2001) and Goldman and Pust (1998) consider a number of ways that the intuitions relevant to philosophical practice might be delimited. Such methods may be considered as proto accounts of the circumstances under which PI's are faulty. The thing to note is that, to the extent that these notions are taken to be accounts of error, they contain implicit empirical claims about how PI's works – claims which have received up to now no formal investigation.