INTUITION AND INQUIRY

by

JENNIFER NADO

A dissertation submitted to the

Graduate School-New Brunswick

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Philosophy

Written under the direction of Stephen Stich

And approved by

________________________________

________________________________

________________________________

_____________________________

New Brunswick, New Jersey

October, 2011
ABSTRACT OF THE DISSERTATION

Intuition and Inquiry

By JENNIFER NADO

Dissertation Director:

Stephen Stich

My dissertation examines prominent arguments for and against the use of intuition in philosophical theorizing. Many of the concerns I raise involve areas of oversimplification - particularly concerning the relationship between the reliability of our intuitions and their evidential status. Specifically, I argue that there are two primary barriers to framing the intuition debate as a simple question about whether intuitions are either unreliable and therefore wholly unsuitable for use in philosophy, or reliable and therefore always suitable for philosophical use. The first of these barriers involves the widespread assumption that intuition forms a fairly cohesive, fairly natural mental kind which can be usefully evaluated for reliability. The second involves the assumption that anyone who questions the suitability of intuitions as evidence must deny that intuitions are reliable enough to constitute knowledge. A further theme of my dissertation involves the relationship between the categories delineated by our intuitive classifications and the categories which serve as the targets of philosophical investigation. I examine two major approaches to this relationship - the ‘mentalist’ view, according to which the targets of philosophical analysis are our own concepts, and the ‘extramentalist’ view, according to which the targets of philosophy are external phenomena ‘in the world’. I
conclude that both projects are legitimate, but that only the former justifies the commonly
held view that there exists a sort of ‘constitutive’ relationship between intuitions and
truth. Ultimately, I advocate a position according to which use of intuition in philosophy
should be restricted, but not eliminated.
ACKNOWLEDGEMENTS

I would like to thank, first and foremost, Dr. Stephen Stich for his invaluable input and advice throughout the entirety of my graduate career. Without his endless patience, his thorough, careful, and insightful criticism, and his constant guidance and encouragement, this dissertation would not have come to fruition.

I would also like to thank the members of my committee, Dr. Alvin Goldman, Dr. Ernest Sosa, and Dr. Michael Devitt. Each of these individuals has brought his own unique perspective and contribution to this dissertation, both in the published work that forms the foundation upon which my research hopes to build, and in the personal involvement each has had in shaping the arguments found in the text. I am deeply grateful for their time and their wisdom.

In addition to my committee members, I am fortunate to have had the input of many other wonderful philosophers during the course of writing this dissertation. In particular, I would like to thank Brian Weatherson, Jonathan Weinberg, Joshua Knobe, Joshua Alexander, Barry Loewer, Martin Lin, Richard Dub, Karen Shanton, Karen Lewis, Matt Benton, Pavel Davydov, Andrew Sepielli, Meghan Sullivan, Allison Hepola, Gabe Greenberg, Alex Morgan, Jen Wang, Ben Levenstein, Marcello Antosh, Leesun Choi, Jonathan Winterbottom, Nick Beckstead, Heather Demarest, and Evan Williams.

To my husband, Michael Johnson, I owe a double debt. First, I am grateful for the many hours we spent in discussion, helping me work through difficult portions of argumentation or fleshing out my understanding of underlying issues. And second, I am
grateful for his endless, unwavering emotional support during my most discouraged and frustrated hours. I am indescribably lucky to have him by my side.

Finally, I would like to thank my parents and sister, whose unconditional love, faith, and support has been a constant treasure, not just through the writing of this dissertation, but throughout my entire life.
# TABLE OF CONTENTS

ABSTRACT OF THE DISSERTATION...........................................................................................................ii

ACKNOWLEDGEMENTS...............................................................................................................................iv

TABLE OF CONTENTS................................................................................................................................vi

CHAPTER 1: THE ROLE OF INTUITION........................................................................................................1

1. Anti-intuitionist arguments..................................................................................................................2
   1.1 The variation argument....................................................................................................................3
   1.2 The calibration argument..................................................................................................................8
   1.3 The restriction argument................................................................................................................14

2. Traditionalist arguments......................................................................................................................18
   2.1 The self-defeat argument................................................................................................................19
   2.2 The parity argument.......................................................................................................................26
   2.3 The constitutivity argument............................................................................................................31

3. Desiderata for the anti-intuitionist......................................................................................................38

CHAPTER 2: WHY INTUITION?..................................................................................................................43

1. Current perspectives on intuition........................................................................................................44

2. A ‘generality problem’ for intuition?.................................................................................................51

3. Evidence for a ‘heterogeneous faculties’ view....................................................................................54

4. Consequences for current debates....................................................................................................67

CHAPTER 3: INTUITION, EVIDENCE, AND THE THREAT OF SKEPTICISM......................................74

1. Evidence in philosophical theorizing.................................................................................................75

2. Intuition and concept application.......................................................................................................82
3. Variation arguments and massive error.......................................................89
4. Philosophy versus everyday cognition..................................................92

CHAPTER 4: DOES INTUITION SUFFICE?......................................................100
   1. The argument from evidence.............................................................101
   2. The argument from concepts...........................................................115

CHAPTER 5: TRUTH AND THEORETICAL UTILITY.................................135
   1. Classical descriptivism and the reliability of intuition........................136
   2. The two-dimensional approach to constitutivity.............................140
   3. The question of theoretical utility.....................................................146
   4. The extra-mentalist approach to philosophical theory......................152
   5. The psychological approach to philosophical theory.......................158

EPILOGUE..................................................................................................164

REFERENCES..........................................................................................168
Chapter 1

The Role of Intuition

Within the past decade, the use of intuition in philosophical theorizing has come under increasing scrutiny. Traditionally, intuition-based methods such as conceptual analysis have formed the core of the philosophical toolbox; however, some philosophers have begun to question traditional practice, raising doubts about the widely-held assumption that intuitions constitute evidence for or against philosophical hypotheses. Many within this minority have argued that appeals to intuition are fundamentally misguided, and that philosophical methodology should be radically revised. Proponents of the conventional intuition-based methodology have responded to these challenges with a variety of arguments, both attacking the anti-intuitionist’s criticisms and defending intuition’s epistemological credentials.

Though the details of particular approaches vary widely, several broad argument types can be identified on each side of the debate. The aim of this first chapter will be to survey the primary argument types present in the literature over the past ten to fifteen years, to identify particular examples of their use, and to provide a cursory assessment of their effectiveness. As my sympathies lie primarily with the anti-intuitionists, a further goal will be to investigate possible lines of response to each of the main defenses of intuition. The end result will not be a refutation of the traditionalist, but rather a

---

1 The taxonomy I propose in this paper is far from exhaustive; I do, however, hope to have covered the most prominent views in the recent literature. There are a few important responses to specific anti-intuitionist arguments which do not fit into my proposed taxonomy. I have not, for example, proposed an argument type for Sosa’s suggestion that the findings of Weinberg, Nichols, and Stich (2001) can be explained by a difference in the propositions the subjects are entertaining rather than by disagreement over a particular proposition (see Sosa (2009)). Specific responses of this sort will instead be discussed during exposition of the arguments they respond to, where appropriate.
suggested research program for the anti-intuitionist – a set of desiderata for a successful account of the evidential failings of intuition. The remainder of the dissertation will consist in a set of preliminary steps along the path laid out by this research agenda.

1. Anti-intuitionist arguments

This first section of the chapter will focus on the arguments of the anti-intuitionists. Critics of intuition have, in recent years, offered two primary styles of attack. The first centers on the phenomenon of variation in intuition, both between and within subjects. This approach has been pursued primarily by members of the experimental philosophy movement, including Stephen Stich, Jonathan Weinberg, and Edouard Machery. The second major anti-intuition strategy is to allege that intuition is illegitimate as a source of evidence because it is not amenable to calibration. This type of argument was first championed by Robert Cummins, but has also recently been defended in a modified form by Jonathan Weinberg.

Finally, there also exists an intermediate camp of philosophers who support intuition’s evidential status to some degree, but who also argue for substantial restriction in its use. It will be useful for the purposes of this chapter to group this sort of argument together with the two anti-intuition arguments described above; for, at the very least, the restrictionist position calls for significant revision of our current philosophical practices. Prototypical restrictionists include Michael Devitt, Hilary Kornblith, and Brian Weatherson.
1.1. The variation argument

Variation arguments share the following very simple argumentative structure. First, it is claimed that if intuitions about some philosophical notion N vary as a function of a certain feature F, then the project of using such intuitions to characterize the nature of N is misguided. Second, evidence is offered in support of the claim that intuitions do in fact vary as a function of feature F. It is then concluded that the project of using intuitions to characterize the nature of N is misguided.

Perhaps the most common feature F is cultural background of the subject. In the paradigmatic study of this kind, Weinberg, Nichols and Stich (2001) presented subjects of either Western or East Asian background with vignettes describing a Gettier scenario; subjects were then asked to judge whether the character in the scenario “really knows” or “only believes”. Though the Western subjects mirrored Western philosophers in judging that Gettier cases were not cases of knowledge, East Asian subjects tended to claim that the Gettier cases were in fact cases of knowledge. Similar results were found with “truetemp” cases – Western subjects’ intuitions mirrored the intuitions of Western philosophers, but East Asian subjects’ intuitions did not. Unless we are willing to embrace epistemic relativism, it seems that we must reject one group’s intuitions as false; however, in the absence of some sort of error theory, it would appear to be nothing but naïve ethnocentrism to doggedly hold that the Western responses are the correct ones. Weinberg et al. conclude that intuition should not be the basis for normative epistemological claims.
Machery, Mallon, Nichols and Stich (2004) found evidence for similar cross-cultural variation in intuitions about reference. Subjects were presented with a version of Kripke’s “Gödel/Schmidt” thought experiment, in which a man named “Schmidt” turns out to have been the real discoverer of the incompleteness of arithmetic - the man called “Gödel” merely stole the proofs and published them as his own. Kripke’s observation was that, in such a scenario, it is intuitive that the name “Gödel” refers to the man we had been calling “Gödel” – contra the predictions of the description theory of names. While the intuitions of Western subjects tended to mirror Kripke’s, thus supporting causal-historical views of reference, East Asian responses to the Gödel/Schmidt case were markedly more descriptivist. Again, Machery et al. take it that there is no reason to assume that Westerners have better intuitions; thus, they claim, considerable doubt is cast on arguments which defend a particular account of reference by appeal to intuitions.2

Cultural variation arguments are but one species of the “variation” argument type; variation arguments can also be constructed by substituting other features for F in the argument schema described above. In order for the first premise of the argument to be plausible, of course, the feature F which produces variation must be irrelevant, in the following sense: F must be such that its variation in the cases under evaluation does not plausibly imply variation in the truth value of the hypothesis at hand.3 Imagine that a study shows that intuitions about whether a case counts as knowledge vary as a function of whether the person in the described case is in possession of a reliable belief forming

---

2 This study in particular has been subject to some fairly convincing methodological objections, particularly the claim that the study failed to distinguish between speaker reference and semantic reference. However, I maintain that the general case for problematic variation in intuition is still quite strong despite the flaws of this study, given the preponderance of less problematic variation studies. I also maintain that the study of reference is not in principle immune to criticisms regarding its reliance on philosophers’ intuitions as the primary source of data.

3 Where “plausibly”, of course, should not be read as “intuitively” – the intended sense is something like “non-negligible possibility given our current theoretical commitments”.

mechanism. No variation argument looms here, for the feature F causing the variation in intuition – possession of a reliable belief forming mechanism – is relevant. It is plausible that the presence of this feature could affect the truth value of an ascription of knowledge.

On the other hand, imagine a study shows that intuitions about whether a case counts as knowledge vary as a function of the order in which cases are presented. A variation argument is now on the horizon, for case order is arguably irrelevant. If we varying nothing but the case order, and yet our intuitions vary, then we have evidence that at least some of those intuitions are false. Swain, Alexander, and Weinberg (2008) have, in fact, run just such a study; they found that subjects were much less willing to ascribe knowledge to the subject in a Truetemp case when they were first presented with a clear case of knowledge. Conversely, subjects were less likely to ascribe knowledge when they were first presented with a clear case of non-knowledge. 4

The more our intuitions are affected by irrelevant features of this sort, the less accurate they must be at tracking the truth. The strength and generality of the variation argument thus increases when one notes that, even outside of the philosophical intuition literature, numerous studies have demonstrated inappropriate variation on various cognitive tasks. To take just one example, Tversky and Kahneman (1981) famously demonstrated the strength of framing effects with their “Asian disease” problem.

4 One obvious response to these findings is to claim that case order is relevant on a contextualist account of knowledge attribution; presenting a clear case of knowledge before presenting the Truetemp case changes the conversational context, such that the subject’s attribution of knowledge expresses a different proposition than the proposition expressed by those subjects who attributed knowledge after first being presented with a clear case of non-knowledge. Swain et al. give two responses to this suggestion. First, even subjects in the same conditions (same case order) gave widely varying responses – from (1) all the way through (5) on a five point Likert Scale. To accommodate such findings, the contextualist must propose a fairly wide notion of context, such that subjects’ encounters prior to the study count towards fixing the context in which they are making their judgments. If we must individuate contexts that widely, it is unclear how we could ever determine that two subjects were in the same context. Second, Swain et al. found no order effects for Fake Barn cases – it is unclear how the contextualist would provide a non-ad hoc explanation for the difference in stability between the two sorts of example.
Subjects were presented with two hypothetical programs for combating an imminent epidemic in which 600 people were expected to die. If the first program is enacted, 200 people will be saved; if the second program is enacted, there is a 1/3 probability that all 600 will be saved and a 2/3 probability that no one will be saved. Subjects showed a strong preference for the former option. However, preferences were reversed when the very same options were reframed – when the outcomes of the programs were expressed in terms of the number of people who would die, rather than the number of people that would be saved. Subjects’ responses in these cases were thus shown to vary as a function of an irrelevant factor, i.e. the particular way in which the scenario is described. Other examples of problematic variation can be found throughout the rationality literature.

Though the data demonstrating variation in intuitions is quite prevalent, there remains some disagreement over its significance. Sosa (2009), for instance, has argued that the variation in Gettier intuitions observed in Weinberg et al. (2001) might be explained by a tendency for the two cultural groups to interpret the vignettes in different ways.

“It is not clear exactly what question the subjects disagree about. In each case, the question would be of the form: “Would anyone who satisfied condition C with regard to proposition <p> know that p or only believe it?” It is hearing or reading a description of the example that enables the subjects to fill in the relevant C and <p>. But can we be sure that they end up with exactly the same C and <p>?“ (Sosa 2009, 107).

It is possible that the cultural differences between the subjects lead them to fill in details not explicitly specified in the vignettes in different ways – perhaps the two groups vary in some crucial background beliefs regarding fake barns, Ford cars, and the like.

This suggestion is of course an empirical hypothesis, but it is not without plausibility. The cultural variation studies performed to date have not attempted to test for the presence of differing background assumptions. This is a genuine weakness in the
variation argument. It is, however, one that might be remedied by careful empirical work in the future. In the meantime, the amassed variation studies should still cause the traditionalist a fair bit of concern. It is also worth mentioning that the ‘differing background beliefs’ hypothesis is much less plausible for explaining away framing effect and order effect findings.\textsuperscript{5}

Sosa provides a second, similar concern for the Weinberg et al. findings. It is possible that the Western and East Asian groups have a purely verbal disagreement; “knowledge” might express different concepts for the two groups. There is no genuine disagreement, after all, if Westerners ascribe knowledge\textsubscript{1} to a subject in a Gettier case while East Asians withhold an attribution of knowledge\textsubscript{2}.

The trouble with this suggestion, however, is that it provides us with no obvious means for answering the normative question – the question of which epistemic states we \textit{ought} to pursue – without admitting some degree of epistemic relativism. Many epistemologists take knowledge to be the primary goal of epistemic activity; if the above ‘divergent concept’ case obtains, however, our tendency to value knowledge\textsubscript{1} over the East Asian knowledge\textsubscript{2} might amount to nothing more then cultural preference.\textsuperscript{6} Sosa appears to be untroubled by this consequence; he argues that both concepts might express valuable epistemic goals, and that our preference for the one should not preclude our valuing the other.

Stich (2009) has rightly noted that this response becomes much less plausible when one considers moral disagreement – we should in no way feel inclined to value, for

\textsuperscript{5} Sosa has a separate reply to the Swain et al. order effect findings. It is a reply which draws an analogy between intuition and perception, and is a good example of what I call a ‘parity’ argument. Parity arguments will be discussed in section 2.2.

\textsuperscript{6} Of course, it may well be the case that knowledge\textsubscript{1} is a better epistemic goal than knowledge\textsubscript{2}; the point is that this thesis can not be defended solely via intuitions about what falls under the term ‘knowledge’.
instance, the moral views of the infamously violent Yanomamö tribe. I would add that Sosa’s response is also problematic for non-normative, philosophically relevant phenomena like belief, time, causation, and the like. If “causation” picks out causation\(_1\) for Westerners and causation\(_2\) for East Asians, we cannot simply embrace cultural relativism about causation - we are instead faced with the task of determining which notion deserves a place in a satisfactory metaphysics. The answer, of course, might be both – say, if the notions play different explanatory roles. Causation\(_1\) and causation\(_2\) might, for instance, stand to “causation” in something like the way nomological necessity and logical necessity stand to “necessity”. But it is highly doubtful that every case of disagreement could be defused in such a manner.

1.2. The calibration argument

It is frequently suggested that the role intuition plays in philosophical inquiry is analogous to the role that observation plays in the sciences. In the sciences, great care is taken to ensure that the observational procedures employed are accurate. A typical way to do this is through calibration. Cummins (1998) uses the following example.

“When Galileo pointed his newly devised telescope at the moon and saw mountains – earthlike blemishes on what should have been a perfect celestial object – it was legitimate for the opposition to inquire whether the apparent mountains were artifacts. The proper response was to point the telescope at something of known size, shape, distance, color and so on to determine what distortions it introduced; to calibrate it, in short” (Cummins 1998, 116-117).

Before we deem a procedure to be an acceptable source of evidence, we ought to apply that procedure to something whose properties are already known – something to which we have independent access. If the output of the new procedure matches our independent knowledge, this bodes well for the procedure’s accuracy.
Unfortunately, according to Cummins, philosophers do not attempt to calibrate intuition. This is no fault of the philosophers; intuition is, in many cases, impossible to calibrate. We simply do not have independent access to the sorts of facts that intuition purportedly reveals. Imagine we are trying to calibrate our moral intuitions; we need to test our intuition against cases involving actions which are independently known to be either morally good or morally bad. But how, other than via intuition, can we determine that a certain act is morally good? The tempting response is that we infer it from widespread agreement; the cases we calibrate our intuition on are the uncontroversial cases. This, however, would be like trying to calibrate a telescope by comparing the results it gave to those given by other telescopes of the same make. Telescopes of that type might all be subject to the same design flaw; this is why independent access to the target is necessary. But it is unclear how independent, non-intuitive access to moral facts could be achieved.

Cummins does suggest that there are some cases in which intuition can be calibrated; we can, for instance, ask what notion of time our best physical theory demands and then check intuition’s deliverances against this notion. Thus, intuition can be calibrated by theory. The catch-22 for the traditionalist, according to Cummins, is that if we were in possession of a well-developed theory which afforded us independent access to the targets of intuition, we would no longer have need for intuition. “Philosophical theory in such good shape is ready to bid the Socratic midwife farewell and strike out on its own in some other department” (Cummins 1998, 118). One might disagree with Cummins on this last point; regardless, the ability to partially calibrate intuition by theory offers cold comfort. In those cases where we have been able to check
our more theoretical intuitions against our best theories, intuition has a fairly spotty track record; stock examples of errors in such intuitions include the naïve comprehension principle in set theory and Kant’s claim that space is necessarily Euclidian.

One response to Cummins’ argument, offered by Goldman (2007), is to claim that calibration against independently validated procedures is too strict a requirement on sources of evidence. Basic evidential sources like perception, memory, and introspection are, Goldman claims, resistant to this sort of calibration; we do not have procedures for accessing the relevant facts that do not ultimately rely on the faculties being tested. Yet we do not reject any of these as sources of evidence. One might mention that Weatherson (2003) and Weinberg (2007) both note that we can, for instance, calibrate one perceptual modality with another – we can use touch to calibrate vision, and so on.\(^7\)

And it does seem that we must have some method for determining the accuracy of our basic evidential sources, for we all agree that there are some contexts in which perception, memory, etc. systematically err.

Nonetheless, it is reasonable to maintain, with Goldman, that we are justified in employing a basic source of evidence even if we have not performed a thorough assessment of its reliability. This is not, however, to say that such sources are immune to criticism. Goldman maintains that a weaker condition holds – we should not be justified in believing that the evidential source in question is unreliable. As mentioned above, there are a number of cases in which intuition has proven to be in error. But of course, we know that perception is fallible as well. Perception’s fallibility does not impugn it as

\(^7\) This line, of course, doesn’t work for every type of perceptual data – we cannot calibrate tastes in this way.
an evidential source; are there reasons to think that the situation will be different for intuition?

Weinberg (2007) and Weinberg et al. (under review) argue that there are. We know quite a bit about the circumstances under which perception is liable to error, and we are quite good at restricting our use appropriately – we do not, for instance, put much stock in visual perceptions in dark rooms. This is not necessarily a result of calibration in Cummins’ sense; there are ways to test the accuracy of an evidential source which do not depend on comparison with the deliverances of independently validated procedures. In cases where we cannot make use of Cummins’ form of calibration, which Weinberg et al. term ‘domain-oriented’ calibration, we may still be able to employ what Weinberg et al. call ‘instrument-oriented’ calibration. We employ instrument-oriented calibration by turning to an examination of the outputs of the evidential source itself.

By examining the output of a new observational device both over time and in varied situations, we may be able to detect problematic functioning. One test is that of consistency; if a microscope pointed at a given object produces inconsistent readings at different times, we then have reason to doubt its accuracy. Another way to test for error is to look for features in the environment which might be illegitimately affecting the device. Weinberg et al. offer the example of a dance troupe in the next room interfering with the readings of a seismic sensor. If we do identify some problematic functioning via one of these methods, the device may be ‘rehabilitated’ – i.e., outfitted with some means of resisting the interfering factor(s). If rehabilitation is impossible, the device may instead be ‘restricted’ – i.e., we may decide to avoid using the instrument under the problematic conditions. If we are not able to successfully calibrate and subsequently
rehabilitate or restrict, either via domain-oriented or instrument-oriented calibration, an
epistemic source whose reliability is in doubt will be unable to regain its credibility.

Weinberg suggests in his (2007) that the real trouble for epistemic sources is not
*mere* fallibility. Trouble arises only when an epistemic source suffers from *unmitigated*
fallibility, or ‘hopelessness’. There are at least four basic sources of epistemic ‘hope’, the
possession of which mitigates fallibility. The first source is external corroboration ala
Cummins; the three others are species of instrument-oriented calibration. The second
source of epistemic hope is internal coherence – in the case of intuitions, coherence both
within and across subjects. The third source of epistemic hope is detectability of
margins; in order for our practice of using a given instrument or procedure as a source of
evidence to be legitimate, we must have some means of identifying the conditions under
which the instrument or procedure is likely to err. Often this ability comes from the
instrument or procedure itself; for instance, vision reveals to us its margins when, e.g.,
dim lighting results in a dim visual image. The final source of epistemic hope is
theoretical illumination – if we have a good theory of the workings of the instrument or
procedure, explaining why it works and what has gone wrong when it doesn’t, then this
can serve to mitigate the epistemic impact of fallibility.

Weinberg argues that intuition appears to lack much by way of epistemic hope.
External corroboration is not forthcoming in domains like ethics; in domains where we
are able to check intuition against theory, intuition often fares poorly. As for internal
coherence, the variation studies discussed in the previous section provide substantial
cause for doubt in many cases, and for an even greater number of cases we simply have
insufficient data. With regard to the third source of hope, intuition *may* occasionally
provide the means for detection of its own margins; intuitions can be felt more or less strongly, and in the case of the more far-fetched thought experiments our intuitions are often weak. On the other hand, strength of intuition is rarely exploited by philosophical practice – we do not, for instance, have standards for measuring and reporting intuitive strength. In addition, strongly intuitive assertions like that of the naïve comprehension principle are quite often mistaken. Finally, our degree of theoretical illumination with regard to intuition is minimal. We simply have very little understanding of the causal routes through which intuition operates. Our ability to calibrate and rehabilitate/restrict our intuitions appears, currently, to be quite low.

Some of these epistemic failings may be remediable; for instance, we may eventually formulate a theory of intuition which is as rich and explanatory as is our current understanding of the workings of vision. It is an open question, however, whether such a theory would support or undermine our confidence in our intuitions. We cannot yet successfully calibrate intuition; once we can, we may discover that intuition’s flaws are so thorough that a program of rehabilitation/restriction would not be worthwhile.

1.3. The restriction argument

It is possible to hold that intuition is indispensable to philosophical inquiry while simultaneously arguing that its current usage is overly promiscuous. Michael Devitt, Hilary Kornblith, and Brian Weatherson have each suggested that our use of intuition should be restricted rather than eliminated. This is not to imply that the anti-intuitionists
discussed in the previous sections aim to reject intuition across the board; many may be sympathetic to some limited use of intuition in, say, basic deductive arguments. Restrictionists, however, differ from the philosophers already discussed in that they offer positive accounts of the indispensability of intuition. The restrictions they suggest are, in fact, direct consequences of their views on the positive role of intuition. In addition, at least in the case of Weatherson, the degree of restriction suggested may be fairly minor.

For Michael Devitt, the restrictionist position flows quite naturally from his view that intuitions are simply a species of theory-laden, empirical judgments – as opposed to the more traditional view that intuitions provide a special sort of a priori insight.

Intuitions, according to Devitt (2006), differ from other empirical judgments only in that they are made in the absence of conscious reasoning. We can identify two types of intuitions which play a role in empirical inquiry. The first are the sorts of intuitions by which we identify members of a given kind under investigation; our intuitions that this is an F but this is not. Call this ‘basic’ intuition. The second sort of intuition, which we might call ‘rich’ intuition, provides more general judgments about the F’s identified by the basic intuitions; a typical rich intuition would be something like “belief plays a central role in producing action”.

There are two stages to an investigation into the nature of a given kind, whether that investigation is philosophical or scientific. During the first stage, we must identify uncontroversial cases of the kind to be investigated. Often this is done in the absence of any theory of the kind we are interested in; in such a case, basic intuitions are crucial. The best sources of appropriate basic intuitions are those persons who have the most empirical expertise with the kind at hand. In some cases – Devitt uses pains as an
example - this may be ‘the folk’. In cases where some scientific theory is available, however, intuitions of the relevant scientists are preferable. This is in sharp contrast to the standard philosophical view, upon which we must take pains to avoid intuitions that have been ‘contaminated’ by theory. Theory-contamination, on Devitt’s view, is a virtue rather than a vice.

The second stage of investigation, once we have identified samples of the kind in question, is to examine those samples and determine what is “common and peculiar” to them (Devitt 2006, 4). Rich intuitions may, at this stage, be a source of hypotheses; but they are in no way necessary. Further, as with basic intuitions, rich intuitions should be trusted only insofar as they reflect strong empirical expertise with the kind at hand. The best method of investigation is direct, scientific investigation of the kind; and where intuition and experimentation conflict, it is intuition which should be rejected.

Hilary Kornblith’s account of the proper use of intuition in philosophy is quite similar to Devitt’s. Kornblith, like Devitt, rejects the a priori view of intuition and takes the activity of philosophy to be analogous to the investigation of natural kinds in the sciences. The purpose of appealing to intuition in philosophy is to “make salient certain instances of the phenomenon that need to be accounted for… much like the rock collector who gathers samples of some interesting kind of stone for the purpose of figuring out what it is that the samples have in common” (Kornblith 1998, 134). Thus, both Devitt and Kornblith take intuition to be primarily a means by which to demarcate a group of cases that appear to form a kind worthy of investigation. Kornblith also agrees with Devitt that these identification intuitions are theory-laden, and that the influence of background theory (insofar as the background theory is accurate) improves rather than
degrades the trustworthiness of intuition. Finally, Kornblith and Devitt both agree that this initial process of identification produces only a rough estimate of the boundaries of a class, and that further theory will in many cases show that some of the initial judgments were mistaken.

The Devitt/Kornblith approach involves the view that philosophical investigation is essentially parallel to the investigation of natural kinds in the sciences. Indeed, Kornblith explicitly states that he takes knowledge to be a natural kind. This is in direct opposition to the views of philosophers like George Bealer, who holds that philosophy is entirely autonomous from the sciences and that philosophical concepts are ‘semantically stable’ – that, unlike natural kind concepts, the external environment makes no contribution to their meaning (Bealer 1996). The Devitt/Kornblith approach is also in opposition to views like Alvin Goldman’s, upon which the aim of philosophy is the investigation of psychological concepts, rather than the kinds they refer to. These opposing views, and their plausibility in relation to the Devitt/Kornblith approach, will be discussed in section 2.3.

Brian Weatherson (2003) formulates his proposal for the role of intuition in the context of a defense of the justified true belief (JTB) model of knowledge. The JTB theory is widely considered to be inadequate due to the intuitiveness of the claim that Gettier cases, while cases of justified true belief, are not cases of knowledge. But why should intuition trump theory in such a case? Weatherson argues that there are four criteria upon which to judge the success of a philosophical theory. First, the theory should not have too many counterexamples. “While a theory can be reformist, it cannot be revolutionary” (Weatherson 2003, 6). Second, the theory cannot have too many
undesirable theoretical consequences. To take Weatherson’s example, a successful ethical theory should not imply that conspicuousness of suffering is a morally relevant feature. Third, the analysis proposed by the theory ought to be one upon which the concept analyzed turns out to be theoretically significant; ad hoc analyses are not successful. Finally, the analysis should be simple. Given that a theory might do better than its rivals on two or three of these measures while doing worse with regard to counterexamples, it seems that there should be at least some cases where theory trumps intuition.

According to Weatherson, these theoretical constraints emerge from the nature of meaning. Weatherson adopts a Lewisian approach, upon which the meaning of a predicate is the property that provides the best balance between validation of our pre-theoretic beliefs and ‘naturalness’. The first two theoretical criteria discussed above test for the former, while criteria three and four test for the latter. Given this view on meaning, it seems plausible that ‘knowledge’ might mean justified true belief – though this property falsifies a few of our pre-theoretical beliefs, there simply aren’t any other reasonably natural properties in the vicinity that would do better.

Weatherson’s view is more conservative than the Devitt/Kornblith view, in an important way. For Devitt and Kornblith, intuitions primarily serve to put forth a plausible ‘first draft’ of a candidate class for empirical investigation. For Weatherson, however, intuitions are an important source of data for determining the meaning of a given term; a meaning assignment cannot be true if it does not accord with a good portion of our intuitions. Meaning facts, in turn, are genuinely relevant to philosophical inquiry. It is sometimes asserted that philosophers are interested in knowledge, and not the term
‘knowledge’. But Weatherson notes that, given the truism that ‘knowledge’ refers to knowledge, if one finds out that ‘knowledge’ refers to justified true belief then one can validly infer that knowledge is justified true belief. Facts about meaning can genuinely inform an investigation of knowledge. Conceptual analysis is therefore retained as a legitimate goal of philosophical activity; the role of intuition in conceptual analysis, however, is somewhat restricted.

2. Traditionalist arguments

This second section of the chapter will focus on supporters of intuition, or traditionalists, who have offered three primary lines of defense for conventional methodology. The first approach alleges that arguments against the evidential status of intuition inevitably end in self-defeat; they undermine the very premises they depend on. Both George Bealer and Joel Pust have offered arguments to this effect.

The second type of defense involves the claim that there is an epistemological parity between intuition and some other, uncontroversially evidential faculty – and that there is therefore no principled basis upon which to reject the one and admit the other. Ernest Sosa has championed a parity with vision; Timothy Williamson has urged a similar position with regard to mundane counterfactual reasoning and concept application.8

The final strategy commonly used by the traditionalist is to argue that intuitions must necessarily provide evidence about the concepts or meanings they express, due to

---

8 Williamson’s position is that there is not just a parity between intuition and other cognitive capacities, but an identity – intuition is not a separate mental kind. Nonetheless, the overall thrust of the argument is the same, as will be shown in section 2.2.
the very nature of concepts or meanings; to use Alvin Goldman’s language, it is in some sense *constitutive* of possessing a concept or understanding a meaning that one has intuitions which reveal truths about the relevant subject matter. This final strategy is undoubtedly the most diverse in implementation. Goldman’s use of it reflects an avowedly psychological take on the aims of philosophy; in a less naturalistic vein, the general approach has also been used by George Bealer and by Frank Jackson.

2.1. *The self-defeat argument*

Self-defeat arguments aim to show that the anti-intuitionist thesis undermines itself; anti-intuitionists argue that intuition should be removed from our evidential resources, but the premises of the anti-intuitionist’s arguments can only be defended by appeal to intuition. One recent well-known instantiation of the self-defeat argument type can be found in Pust (2001). Pust’s target is the ‘explanationist’ objection to the evidential status of intuition, one statement of which can be found in Harman (1977). Harman claims that the best causal explanation for our having the moral intuitions we do does not advert to the truth of the intuited moral propositions. Because of this, we are not justified in using moral intuitions as evidence for the truth of moral facts. This, Pust claims, reflects a general ‘explanationist’ criterion of justification – one is justified in believing only those propositions which either a) report the occurrence of judgments or

---

9 I have not included ‘explanationist’ arguments in my taxonomy of the primary arguments against intuition, simply because explanationist arguments have not been prevalent in the intuition literature over the last fifteen or so years. However, Weinberg’s ‘theoretical illumination’ criterion for hopefulness perhaps reflects something of the spirit of the objection – the natural reading of his view is that philosophical use of intuition will be (partially) validated if it turns out that there is a causal pathway between intuitions and the facts they purport to reveal, paralleling the case of vision, providing an explanation of why we have the intuitions that we do – and, further, if that explanation makes plausible the hypothesis that intuition is reliable.
observations, or b) figure in the best explanation of the occurrence of those judgments or observations.

Pust argues that Goldman’s (1992) argument against metaphysical intuitions embraces the explanationist criterion of justification as well – Goldman writes that we should only be confident that our intuitions reflect metaphysical fact if we have an account of their psychological nature which explains how they could reliably indicate metaphysical truths. The general argument form that Pust takes both Harman and Goldman to endorse runs as follows:

(1) Aside from propositions describing the occurrence of her judgments, S is justified in believing only those propositions which are part of the best explanation of S’s making the judgments that she makes.

(2) Propositions regarding domain D are not part of the best explanation of S’s making the judgments that she makes.

(3) S is not justified in believing any propositions regarding domain D.

The first premise is the explanationist criterion of justification. The domain D in the second premise is to be replaced with the particular domain one is questioning. Thus, for Harman, premise two is that moral propositions are not part of the best explanation of the subjects’ judgments; for Goldman it is that metaphysical propositions (about some topic X, perhaps) are not part of the best explanation of the subjects’ judgments. The conclusion states that we are not justified in believing any proposition regarding the problematic domain.
Pust notes that this argument schema is potentially self-defeating, however, as can be seen by substituting epistemology for domain D. The conclusion then becomes

\((3^*)\) S is not justified in believing any epistemological propositions.

But premise (1) is clearly an epistemological proposition; thus, the conclusion undermines our justification in believing the premise. For this claim of self-defeat to have general application, of course, we must add a further premise: Pust must hold that the epistemological version of the explanationist argument should be endorsed by anyone who endorses any version of the explanationist argument. There is no inconsistency in holding that moral facts do not factor into the explanation of moral intuitings, but that epistemological facts do factor into the explanation of epistemological intuitings. Pust argues, however, that “propositions about epistemic justification seem, as a class, as likely to fail the requirement of premise (1) as the other classes of propositions attacked by the explanationists” (Pust 2001, 250).

This last is to my mind debatable, as I will argue in a later chapter; however, for the moment, let us grant it for the sake of argument. Pust’s self-defeat objection is nonetheless unconvincing, at least against Goldman’s argument and against the anti-intuition arguments discussed in the previous section. For the forms of these arguments are just not plausibly captured by Pust’s explanationist argument schema.

Recall the conclusion in Pust’s argument schema: S is not justified in believing any propositions regarding domain D. As stated, the conclusion of the argument is clearly far too broad; perhaps Harman intended to motivate the threat of moral
skepticism, but surely Goldman thinks that we can come to know metaphysical propositions - possibly even some of the very same metaphysical propositions we initially intuited. If Goldman did not think this, his discussion of prescriptive metaphysics in the very paper Pust cites would simply be nonsensical. Instead, Goldman’s objection is that, if our best account of metaphysical intuition does not imply that metaphysical intuition reliably indicates truth, then we need to get at metaphysical truths in some other way. The conclusions of the anti-intuitionists are presumably similar - intuition does not reliably reveal philosophical facts, but this merely means that we must revise our methodology, not that philosophical progress is impossible.

Perhaps we can reformulate Pust’s presentation of the explanationist argument in the following manner:

(1) A subject S is justified in believing a proposition P on the basis of an intuition that P only if the best explanation of S’s having that intuition invokes the truth of P.
(2) For any proposition P about some domain D, the best explanation of S’s intuiting that P does not invoke the truth of P.
(3) For any proposition P about domain D, S is not justified in believing that P on the basis of an intuition that P.

Given this more plausible formulation, does Pust’s self-defeat argument against the explanationist go through? It clearly does not. When we substitute epistemology for domain D, the conclusion simply requires that one’s justification for the epistemological claim (1) not rest on an intuition that (1). But there are all sorts of ways in which one
could be justified in believing (1); notably, one could be justified in believing it because it follows from one’s best epistemological theory. And belief in one’s best epistemological theory could be justified because it explains all sorts of things, be they intuitions or otherwise.

It often seems as though Pust’s real worry is that, at the end of the day, we will have to make use of intuition somewhere. In fact, Pust suggests that the justification for belief in the original argument’s premise (1) – that is to say, the explanationist’s epistemological principle – can only be found in intuition. More generally, Pust claims that “intuitions seem necessary for the justification of epistemological claims and so seem necessary for the reflective construction of any epistemological doctrine” (Pust 2001, 251). A similar concern has, in fact, been expressed in somewhat more detail in an earlier self-defeat argument found in Bealer (1992). Bealer’s self-defeat argument aimed not at recent attacks on intuition, but at radical empiricism – we will see if it might be made plausible for use against contemporary anti-intuitionists.

The empiricist, Bealer writes, holds that one’s evidence consists only of one’s observations or experiences – in other words, she wishes to reject intuition as a source of evidence. To use Bealer’s terminology, she wishes to formulate an intuition-free alternative to our “standard justificatory procedure” (which, Bealer claims, endorses use of intuition). Bealer objects that the empiricist, in formulating this alternative procedure, violates its ban on intuition-based inquiry. The empiricist must surely make use of basic epistemic terms like ‘observation’, ‘theory’, and ‘explanation’ in formulating her new procedure. But how does the empiricist determine what counts as an observation, as a theory, as explanation, or as justification? These basic epistemic classifications – which
Bealer calls “starting points” – are arrived at via intuition, even for the empiricist. The empiricist’s alternative procedure, then, undermines itself.

We’ve noted that there is more than one way to be justified in believing some given proposition. The empiricist/anti-intuitionist might claim that she has non-intuitive justification for making the epistemic classifications that she does; more plausibly, Bealer suggests that she might claim that although she initially formulated her starting points by use of intuition, she no longer relies on intuition for her current justification. Bealer grants this move, and claims that empiricists are caught in a fatal dilemma nonetheless.

Intuitions about starting points are either reliable or they are not. If they are reliable, then they are eligible to serve as evidence and the empiricist’s rejection of them is unwarranted. If they are not, then the starting point judgments that the empiricist initially formulated on the basis of intuition are prone to error. This error, Bealer claims, will be reflected in the theories that result from those starting points – theories which include the empiricist’s epistemological principles. Bealer notes that, during the process developing a theory, it is sometimes possible to identify and correct erroneous judgments – just as we have been able to identify contexts in which our perceptual faculties err. But note that in the case of perception, it is the very fact that our perceptions are on the whole reliable that allows us to identify error successfully. If intuition is unreliable, then error at the theoretical level will not be remediable.

There is a legitimate question for the anti-intuitionist here – given that all theories need some ‘starting point’ classifications, how are we to go about formulating accurate theories if our classification intuitions are unreliable? It’s unclear, however, why Bealer thinks that the anti-intuitionist can’t provide a plausible answer to this question. Why
can’t we employ our other cognitive resources in order to identify and expunge errors generated by our initial intuitions? There is a massive body of propositions which we are justified in believing; it’s plausible that this body is sufficient for the construction of theoretical principles which could lead us to correct errors in our more unreliable classification intuitions. The anti-intuitionist, after all, does not question our use of perceptual observation; in addition, no anti-intuitionist has questioned the use of logic or mathematics.

One might object that our most basic logical and mathematical views, such as our endorsement of *modus ponens*, can be justified only by intuition. Thus, the anti-intuitionist must be relying on intuition after all. Even if the anti-intuitionist were to grant this, however, her views would not be undermined. The reason is this: intuition is not likely to be monolithic. The anti-intuitionist typically aims her barbs at intuitions involving conceptual classification – intuitions that such-and-so case is an instance of knowledge, or of right action. The psychological mechanisms that produce these sorts of intuitions are quite plausibly separate from the mechanisms underlying our use of fundamental logical rules. If this is right, then the reliability of one does not entail the reliability of the other.

Beyond observation, logic, and mathematics, the anti-intuitionist surely needs to admit at least *some* of our classification intuitions; she doesn’t want to question, e.g., my intuition that what I am currently typing on counts as a keyboard. The challenge, then, is to identify an admissible subset of our classification intuitions in a non-ad hoc manner. There are several possibilities here; for instance, Weinberg (2007) suggests that intuition is more likely to err on the more esoteric, non-prototypical cases of a given concept - on
the central cases, we may be fairly reliable. One might also argue that it is only the abstract concepts on which we are likely to err. Finally, one might claim that intuition is on the whole reliable, but that this reliability is not sufficient to justify the centrality it holds in philosophical inquiry. Recall Weatherson’s view – the very nature of reference implies that we will have a fairly high proportion of true intuitions, but one can nonetheless object to the nearly unassailable status philosophers grant to certain counterexamples.

2.2. The parity argument

As noted in section 1.2, the idea that the role of intuitions in philosophy parallels the role of perception in the sciences is fairly attractive. Another tempting thought arises given the uncontroversially justified, ‘basic’ status of perception. Could intuition be similarly basic? How deep are the epistemological similarities between perception and intuition? If it could be shown that intuition’s epistemological properties are similar to perception’s in some relevant ways, one might be able to thereby defend the use of intuition. Since we take perception’s evidential status to be non-negotiable, one might be able to argue by parity that we ought to extend the same status to intuition.

The clearest epistemological similarity between perception and intuition is the following: both are fallible. Further, both exhibit failures which are not merely occasional or random – in many cases, the failures are systematic. This has not prompted us to abandon perception as an evidential source; Sosa (1998, 2007) argues that it should not present a reason to abandon intuition. Studies in experimental philosophy have
shown that our intuitions vary when case order is reversed, or when descriptions are reframed; but, Sosa claims, this is simply analogous to perception’s susceptibility to various errors in unfavorable conditions.

“Surely the effects of priming, framing, and other such contextual factors will affect the epistemic status of intuition in general, only in the sort of way that they affect the epistemic status of perceptual observation in general… the upshot is that we have to be careful in how we use intuition, not that intuition is useless” (Sosa 2007, 9).

We have already discussed, in section 1.2, some of the ways in which intuition might plausibly differ from perception with regard to the epistemological impact of these sorts of errors. Sosa’s suggestion that we simply employ caution only helps if we know what to be careful for; intuition’s poor scores on Weinberg’s four criteria of hopefulness imply that, at present, we are not capable of restricting intuition in appropriate ways. But for the purposes of argument, let us imagine that at some future date we have gained, perhaps via Weinberg’s instrument-oriented calibration, a good deal of insight into the workings of intuition. Further, let us suppose that intuition has fared reasonably well, in a reasonable number of contexts. A question then suggests itself. If intuition is to be treated as analogous with perception, then presumably the role of intuition in philosophy ought to be something like the role of perception in the sciences. Do we philosophers treat intuition in the same way as scientists treat perception?

First, notice that there is a prima facie difficulty in evaluating the analogy we have been pursuing between perception and intuition. Where does intuition end and perception begin? Imagine that I am looking at a cup on a table, and I spontaneously judge “that is a cup”. Do I make this judgment on the basis of perception, or on the basis of intuition? Certainly this is an instance of perception; but isn’t there something like intuition involved in this perception? Paradigm examples of philosophical intuitions
include the judgment that the Gettier case is not a case of knowledge, the judgment that computers are not intelligent, and the judgment that killing a man to distribute his organs is wrong. Each of these is a classification intuition, of the form “x is (not) an F”. When I judge the Gettier case, I judge whether or not it belongs in the knowledge category. When I judge the item before me, I judge whether or not it belongs in the “cup” category. How do these two judgments differ?

Sosa makes a distinction which suggests that my “cup” judgment may be ruled out as a case of intuition, at least in the philosophical sense. “Although perceptual and introspective apprehensions need involve no reasoning, they are not intuitions” (Sosa 1998, 257). Perhaps mental episodes like my cup judgment fall under “perceptual apprehensions”. At the very least, Sosa’s distinction suggests that the parity he argues for is not between intuition and raw sensation, but between intuition and perceptual classification judgments like “this is red” or “this is square”.

With this distinction in mind, I wish to make a few brief observations about the role of perceptual judgments in the sciences. It is beyond the scope of this chapter to fully explore this analogy; my purpose in this survey is only to suggest plausible routes of response, not to pursue them. That said, I suspect that an anti-intuitionist could plausibly make the following claim: if we intend to model our use of intuition on the use scientists make of perception, we’re doing it wrong.

The philosophical investigation of knowledge, for example, has proceeded in something like the following way. We began with a set of data points of the following form: x is a case of knowledge, y is not a case of knowledge, etc. In order to explain this data, we proposed theoretical analyses of knowledge with an aim for providing
something like the “best fit”, capturing as many of the intuition data points as possible. Consider an analogous strategy with vision. Start with a set of data points of the form: *x is red, y is blue*, etc. Propose a theoretical analysis of each color which best captures these data points. There’s something not right about that strategy; it just doesn’t seem like that’s what science does.

Our scientific understanding of red did not proceed via armchair reflection on what the observed items we judged to be red had in common. To be sure, our ability to discriminate colors was indispensable to getting our scientific theories of color off the ground – but at the end of the day, it turned out that the divisions we were making between red, blue, etc. were fairly arbitrary divisions of the light spectrum. We might note that we are still justified in having beliefs of the form “the ball is red”, and that these beliefs are (under normal circumstances) true. But science is interested in carving up the world in a different way. Analogously, perhaps the naïve categorizations our intuitions reveal are not the divisions relevant for our theoretical projects in philosophy.

We have thus far been pursuing an argument for parity between intuition and perception. There is a variant on this argument, proposed by Williamson (2004, 2007), which proceeds in a similar vein. Williamson notes the indispensability and reliability of our counterfactual reasoning and our practices of applying concepts in judgment. Williamson then argues that philosophical intuition, rather than being some *sui generis* mental activity, is simply an application of these basic cognitive capacities. The upshot of this is the same as for the argument for parity with perception – if you grant evidential status to the one, you had better grant evidential status to the other. Anti-intuitionists

---

10 We might also note that the ‘scientific’ technique just proposed for theorizing about redness bears a strong resemblance to the notoriously unsuccessful method of introspectionism in psychology.
cannot consistently reject philosophical intuition while simultaneously embracing other uses of concept application. Worse, if anti-intuitionists bite the bullet and reject concept application across the board, their position becomes one of extreme skepticism.

I am extremely sympathetic with Williamson’s assertion that philosophical intuition is continuous with concept application generally. Thus, rather than attempt to argue against the identification, I suggest that anti-intuitionists should attempt to show that the application of this general cognitive ability is legitimate in some contexts, but not in others. Specifically, that it is legitimate in everyday cognition, and illegitimate when used as the primary source of data for a philosophical theory. As discussed earlier in the context of Bealer’s self-defeat argument, this can be done in a number of ways – one can argue that this capacity is unreliable on outlandish or abstract cases, or that the capacity is reliable enough for everyday cognition while being insufficiently reliable for successful philosophical theorizing. The latter suggestion will be pursued further in a later chapter.

2.3. The constitutivity argument

Constitutivity arguments are characterized by their claim that intuitions must necessarily reveal truths (at least, in suitably good cognitive conditions), due to the existence of some sort of constitutive tie between intuition and meanings or concepts. That intuition is generally reliable simply follows from the existence of this constitutive relation. This is not to suggest infallibilism; it is agreed on all sides that we may still err, if we are inattentive or if we do not reflect appropriately. But these errors must lie at the level of performance, rather than the level of competence.
Bealer’s version of this form of argument involves the assertion that intuition has a ‘strong modal tie’ to the truth; and further, that the existence of this modal tie implies that philosophy is both autonomous from and authoritative over the sciences. It is important to note that Bealer is concerned only with ‘rational’ intuition; rational intuition is to be distinguished from physical intuition - which includes, e.g., my intuition that an unsupported object will fall. Rational intuitions, unlike physical intuitions, present their contents as necessary.

The most plausible argument Bealer offers for the existence of a strong modal tie between rational intuition and the truth is his ‘multigon’ example. Suppose a woman introduces through use (as opposed to via stipulation) a new term, ‘multigon’. She applies this term to pentagons, hexagons, and so forth, but she has, at time $t$, neither applied the term ‘multigon’ to, nor withheld it from, triangles or rectangles. Now suppose that at time $t_1$ she considers whether or not triangles and rectangles are multigons. Suppose further that the following conditions hold – 1) she is suitably intelligent, attentive, and otherwise in high-quality cognitive conditions, 2) her term ‘multigon’ expresses a definite concept, and 3) she determinately possesses this concept. ‘Determinate’ concept possession is a term introduced by Bealer, and is to be contrasted with ‘nominal’ concept possession. We possess a concept nominally if we are able to have propositional attitudes towards propositions whose contents contain that concept. We possess a concept determinately if we possess it nominally, and in addition we do so without misunderstanding or incomplete understanding.

Given the conditions just enumerated, Bealer argues the following. At time $t_1$, the woman will judge that triangles and rectangles are multigons if and only if the property
of being a multigon is identical to the property of being a polygon. The woman will judge that triangles and rectangles are not multigons if and only if the property of being a multigon is identical to the property of being a polygon with five or more sides. Thus, the woman’s intuitions have a strong modal tie to the truth.

As stated, however, the account is unsatisfying. Bealer has shown that, if we possess a concept determinately, then our intuitions with regard to that concept will have a strong modal tie to the truth. By the definition given, this is merely to say that if we possess a concept without misunderstanding or incomplete understanding, our intuitions will necessarily be generally true. On what I take to be the standard interpretation of the word “misunderstanding”, this seems to say that if our possession of a concept involves no falsehood, our intuitions about that concept will necessarily be generally true. But the anti-intuitionist is not going to lose much sleep over the fact that there is a strong modal tie between non-falsity and truth.

For his account to have any bite, Bealer needs to show that we do possess philosophically relevant concepts determinately. A step towards such an argument is given by Bealer’s claim that philosophical terms are ‘semantically stable’ – that knowledge of their conditions of application does not require any contingent knowledge about the speaker’s external environment (in contrast to terms like ‘water’, whose meaning depends on the nature of the watery stuff in one’s environment). Since no empirical knowledge is required in order to possess these concepts determinately, there is no barrier to determinate, a priori philosophical understanding - and, “intuitively, it is at least possible for most of the central concepts of the a priori disciplines to be possessed determinately by some cognitive agent or other” (Bealer 2000, 12).
This move does, in principle, appear to lend some plausibility to the thesis that intuitions possess a constitutive, necessary tie to truth. However, the question of whether philosophical terms do in fact possess semantic stability is still open to debate. As I will argue in a later chapter, it’s not at all clear that philosophical terms are as distinct from natural kind terms as Bealer supposes. It is also worth noting that there has still been no real positive explanation of the ‘strong modal tie’. If Bealer is right, then we do not need knowledge of the contingent features of the external world in order to understand when something counts as, e.g., a case of knowledge. But surely this is insufficient to guarantee the necessary reliability of our intuitions about knowledge. After all, presumably no knowledge of the contingent features of the external world is required in order to come to know the truths of mathematics; but this fact alone does not guarantee that we are reliable at mathematical reasoning. Nor does that fact explain whatever reliability our mathematical capacities may possess (indeed, our access to mathematical truths remains deeply puzzling).

Though Bealer’s account, in my mind, fails to explain the proposed necessary reliability of intuition, such an explanation might be derived from an account of the nature of meaning. A meaning-based approach to the a priori accessibility of philosophical knowledge has, in fact, recently been offered by Frank Jackson in his book *From Metaphysics to Ethics*. To begin, Jackson argues that philosophy requires ‘serious metaphysics’. Serious metaphysicians work with a limited ontology - for the physicalist, this ontology should more or less consist of the properties, objects, relations etc. employed by completed physics. Most philosophically interesting terms will not explicitly occur in the language describing the fundamental ontology. However, the
existence of ‘higher-level’ entities may be entailed by the basic ontology, if the higher-level entities are supervenient on the lower-level entities. Being entailed in this way is both necessary and sufficient for an entity’s inclusion in the serious metaphysician’s implied ontology. Jackson calls this principle ‘entry by entailment’.

For each entity described in a high-level vocabulary, the metaphysician must either ‘locate’ that entity by showing how its existence is entailed by facts described in the physical vocabulary of her basic ontology, or she must eliminate the higher-level entity from her ontology entirely. Here is where conceptual analysis is required; for, according to Jackson, “conceptual analysis is the very business of addressing when and whether a story told in one vocabulary is made true by one told in some allegedly more fundamental vocabulary” (Jackson 1998, p. 28). Conceptual analysis, by examining folk intuitions about the extensions of terms in possible scenarios, reveals the implicit understanding of the folk and uncovers the meanings of their terms, and thus the entities that will have a place in our implied ontology.

This is a genuinely explanatory constitutive approach; the fundamental idea is that the folk possess mental theories or descriptions which determine the meanings of their words. Given that one’s intuitions reflect these theories, then, one’s intuitions reflect facts about the meanings of one’s terms. Further, our intuitions about hypothetical cases will necessarily be by-and-large true. An example will make this clear. Imagine that the mental description I associate with my word “knowledge” is simply “justified true belief”. The extension of “knowledge”, on Jackson’s account, is determined by the

---

11 Jackson’s full account is considerably more complicated than this, due to his use of two-dimensional semantics. The details arising from the two-dimensional account, however, are not relevant for the purposes of this chapter. For those familiar with the two-dimensional framework, read “A-extension” for “extension” in the paragraphs to follow.
mental description I associate with “knowledge” – it is the set of things that satisfy that description. So the extension of “knowledge” is the set of justified true beliefs. Now, whenever I consider a hypothetical scenario in which an agent has a justified true belief, my mental description will under normal circumstances lead me to intuit that that agent has knowledge. Thus, something is knowledge if it fits the mental description I associate with “knowledge”, and I will typically judge something to be knowledge if it fits the mental description I associate with “knowledge”. So, in normal cognitive circumstances, my knowledge attributions will be true – due to the very nature of meaning.

One potential trouble with this version of the constitutivity strategy is, essentially, that the divisions effected by one’s mental descriptions may not be important divisions from a theoretical standpoint – depending on the theorist’s goals, the ‘implied ontology’ Jackson’s method will generate may not be an explanatorily useful one. On descriptivist accounts of reference, a term refers to that which falls under the description associated with it; however, not all descriptions delineate important kinds. It is open to the anti-intuitionist to claim that philosophy should be concerned with phenomena that, for instance, aid in explanation and prediction or factor into laws; and further, that there is no reason to think that folk theory reflects categorizations that play those theoretical roles. On Jackson’s account, intuition reliably produces true beliefs; but on descriptivist accounts of reference, truth comes cheap. In order to defend the philosophical use of intuition, it is not enough to show that intuition reliably tracks truth. One must also show
that the truths revealed by intuition are truths about kinds which play an important role in
accomplishing the theoretical goals one is currently pursuing.12

Alvin Goldman’s version of the constitutivity argument, by contrast, avoids this
problem. His strategy is to claim that the truths revealed by intuition are simply truths
about one’s concepts, in the psychological sense of concept – and further, that a primary
aim of philosophy is to uncover these psychological truths. On Goldman’s account,
possession of a concept is to be understood as possession of a psychological structure
underlying one’s use of a given natural-language term. Given this characterization of
concepts, it seems to simply follow that possession of a concept which underlies a natural
language term “F” will involve a disposition to judge “x is an F” when and only when x
satisfies that concept. What it is to have a given concept is to have a psychological
structure which disposes one to make categorizations in accordance with the content of
that concept.

As with Jackson’s account, this version of the constitutivity approach is genuinely
explanatory; the necessary reliability of intuition is grounded in the very nature of
concept possession. However, this story only really explains why my intuitions are
reliable indicators of facts about my individual, psychological concepts. We may be able
to move from facts about individual concepts to facts about shared concepts or word
meanings, if the members of our community have concepts which are substantially
similar. Where there is substantial disagreement among members of the community,
however, even this may not be possible. Goldman acknowledges this, and his account

12 Jackson has some resources for avoiding this objection; he might, for instance, build the theoretical
importance requirement directly into the mental description. A full discussion of this strategy will be given
in a later chapter, but the general response the anti-intuitionist should make is that this severs the direct link
between intuition and truth in much the same way as does the causal descriptivist strategy of building the
causal source of one’s word into the mental description.
thus expresses a very modest assessment of the role of intuition. On the other hand, the account does imply that intuitions are genuinely evidential, in that they grant reliable access to truths about our personal psychological concepts; further, this evidential role is not undermined by phenomena like cultural variation. Finally, truths about personal psychological concepts are plausibly explanatorily important, particularly if one’s project is avowedly psychological; Goldman will not, therefore, face the problem just raised for Jackson’s “implied ontology”.

Goldman claims that viewing philosophical analysis as targeting psychological concepts provides justification for a good portion of our actual philosophical practices. For instance, it explains why philosophers place a high value on pre-theoretical intuition – if one’s intuition is influenced by explicit theory, it will not reflect one’s underlying concept. I would note, however, that Goldman’s account does not make sense of certain other philosophical practices – for example, ‘biting the bullet’ when one’s theory produces counter-intuitive results. Further, if our aim were purely psychological, undesirable theoretical features like inconsistency or ontological promiscuity would not provide a reason to reject an analysis – after all, we should not assume that our personal psychological concepts avoid contradiction or make appropriate use of Occam’s razor.

One begins to suspect that current philosophical methodology often reflects a running together of both Goldman’s aim of characterizing our concepts as well as a Devitt/Kornblith type aim of delineating explanatorily useful kinds. In fact, it is entirely consistent to pursue both projects – and as such, the anti-intuitionist need not find fault in Goldman’s account of the evidential value of intuition. In some cases, the two approaches may even complement one another; as Goldman notes in earlier writings on
metaphysics, an understanding of psychology can indicate to us that some of our metaphysical distinctions are not objective, but merely reflect innate tendencies to e.g. group perceptual elements together in certain ways (Goldman 1987). Psychological investigation can provide the metaphysical prescriptivist with error theories.

3. **Desiderata for the anti-intuitionist**

Having surveyed some primary forms of argument for and against the use of intuition, and having offered some initial criticisms of the traditionalist arguments, I will now summarize what I see as the central remaining obstacles which a critique of intuition-based methodology must overcome. Attempts to address these obstacles will form the basis of the remainder of my dissertation.

Discussion of the ‘self-defeat’ approach to the defense of intuition led us to the following observation: while such arguments may well show that it is impossible to consistently reject use of intuition across the board, the anti-intuitionist may still have room to critique standard philosophical methodology. If she can justify the claim that intuition is not monolithic – and that, therefore, she can accept some types of intuition (say, logical intuitions) while condemning others (say, moral intuitions) – the force of self-defeat arguments will be greatly diminished. It is my belief that an examination of the actual cognitive mechanisms underlying intuitive judgments will reveal that intuition is, in fact, highly heterogeneous – and that, therefore, its epistemological status cannot be articulated with a broad brush. A preliminary survey of the evidence for heterogeneity will form the topic of chapter 2; I will conclude that it is probable that we have multiple
more-or-less domain specific mechanisms underlying different areas of philosophical cognition.

The self-defeat argument also hints at a further, related question: exactly how far should the anti-intuitionist critique extend? This concern also comes to the fore in Williamson’s version of the parity argument; if intuition is simply a form of concept application, then it is incumbent on the anti-intuitionist to motivate a division between the general practice of concept application we aim to keep and the philosophical use of concept application we aim to reject. If rejecting intuition amounts to rejecting all uses of concept application, the anti-intuitionist position begins to resemble a very radical form of global skepticism.

Note that the ‘heterogeneity’ reply just suggested only partly addresses this worry; even if, say, epistemological cognition forms a proprietary cognitive domain which must be assessed separately from (e.g.) logical cognition, there still remains the worry that a rejection of the use of intuition for epistemological theorizing might require us to abandon use of epistemological categorization in daily life, as well. In chapter 3 I will suggest that this worry is not as pressing as it might seem. It is quite plausible that intuition (across a variety of domains) is on the whole quite reliable – sufficiently reliable, in fact, for ordinary purposes. However, the degree of unreliability that intuition does exhibit might nonetheless lead to substantial error at the theoretical level, thereby undermining its suitability as the primary source of evidence in philosophical theorizing.

Chapters 4 and 5 will tackle constitutivity arguments for the use of intuition. In chapter 4, I will undertake an extended examination of George Bealer’s account of the use of intuition in philosophy. Earlier in this chapter, I suggested that Bealer’s account
provides no positive explanation of the hypothesized ‘strong modal tie’ between intuition and truth. Rather, Bealer takes the existence of such a tie to be necessitated by the success of his self-defeat arguments. Expanding on discussion from chapter 2, I will suggest that this reasoning is flawed, and that the limited success of self-defeat arguments simply cannot suffice to underwrite confidence in a strong modal tie for intuition. I will also undertake an extended discussion of Bealer’s suggestion that philosophical terms are ‘semantically stable’. I will argue that, on the contrary, it is highly plausible that the meanings of philosophical terms are, like the meanings of natural kind terms, dependent on contingent features of the external world.

Chapter 5 will consist in a discussion of both Frank Jackson’s and Alvin Goldman’s versions of the constitutivity argument. As noted earlier in the chapter, Jackson’s account of the role of conceptual analysis invokes facts about meaning to explain the proposed necessary reliability of intuition. This is accomplished through a broadly descriptivist account according to which the implicit theories possessed by an agent determine the meanings of her terms. I will argue that this approach, while guaranteeing that intuition is on the whole reliable, fails to address the question of whether the categories delineated by our intuitive classificatory practices are categories that suit our theoretical purposes.

If our purpose is a descriptive, psychological one, as Goldman urges, then the answer to this question may be yes; however, I will argue that the psychological project at best only fits current philosophical practice very imperfectly. If our purpose is, on the other hand, to characterize theoretically mind-independent phenomena ‘in the world’, then there is no reason to assume that, assuming Jackson’s descriptivist account of
reference, our words refer to theoretically important kinds; thus, there is no reason to assume that true intuitions about a given kind will further our theoretical projects. If an argument of this form can be adequately defended, then the anti-intuitionist will be able to claim that Jackson has not yet provided a plausible method for achieving our philosophical goals in a purely *a priori* fashion.

Ultimately, the primary goal of this dissertation will be to clarify the conceptual territory surrounding arguments for and against the use of intuition. One general theme will be that many aspects of the intuition debate have been oversimplified – for example, it has generally been assumed that intuition is a single, cohesive mental faculty, and it has generally been assumed that a demonstration of the overall reliability of intuition would suffice to demonstrate its suitability for philosophical use. Chapters 2 and 3 will question these assumptions. A second theme is the relationship between intuition and the targets of philosophical analysis – this will be explored through an examination of the semantic properties of philosophical terms in chapter 4, and in chapter 5 through a discussion of the theoretical consequences of viewing the aim of philosophy as characterizing internal, psychological features versus external, ‘objective’ categories in the world.

Though the dissertation will be light on positive recommendations, the picture that will emerge during my critiques will be one of a broadly restrictionist character. Self-defeat arguments and Williamsonian parity arguments should convince us that some uses of ‘intuition’ must be preserved, on pain of skepticism. However, the heterogeneity of intuition and the diversity of epistemological goals to which it may be applied should also convince us that full-fledged traditionalism is not the only alternative; a bounty of intermediate positions are available, and given the plausibility of anti-intuition
arguments, an intermediate position seems most tenable. As to more specific positions on the epistemological worth of various types of intuition for various theoretical projects, to my eyes it is early days yet. More work – much of it empirical – remains to be done before specific prescriptions can be made; nonetheless, I am confident that some revisions to our philosophical methodology are bound to be required. At the very least, however, continued examination of such standard theoretical tools as intuition cannot but improve the philosopher’s understanding of her own discipline.
Chapter 2

Why Intuition?

Recent debates over philosophical methodology have overwhelmingly focused on the epistemological merits of a type of mental state commonly referred to as ‘intuition’. Philosophers on both sides have tended to assume that the defensibility of traditional, \emph{a priori} philosophical method fundamentally depends upon whether or not this mental state can be shown to constitute a respectable evidential source. More often than not, the question is framed in terms of intuition’s \emph{reliability}; opponents of traditional methodology attempt to show that intuition fails to reliably track the truth, while defenders attempt to refute these claims and to show that intuition is sufficiently reliable after all. The very fate of intuition-based inquiry is taken to hang on the outcome of these disputes.

In this chapter I will argue that this entire dialectic is somewhat misguided. The mental states which are generally assumed to fall under the category of ‘intuition’ comprise a highly heterogeneous group; from the point of view of psychology or of neuroscience, in fact, ‘intuitions’ appear to be generated by several fundamentally different sorts of mental processes. If this is correct, then it may be a mistake to focus on the ‘reliability of intuition’. The term ‘intuition’ may simply carve things too broadly; the reliability of one type of intuition may tell us next to nothing about the reliability of other types. Rather than debating the evidential status of intuition as a whole, philosophers interested in methodology might do well to focus their investigations much more narrowly.
The first section of the chapter will examine the propensity of philosophers on both sides of the intuition debate to frame discussion in terms of the reliability of a supposedly cohesive, supposedly natural kind of mental state – intuition. In section two, I will briefly discuss the ‘generality problem’ for reliabilism, arguing that it raises the important possibility that intuition may not pick out a mental state of the appropriate grain to be usefully evaluated for reliability. The generality problem itself, however, will not be my focus – rather than attempting to offer any sort of solution to the generality problem, I will instead argue for a far narrower thesis. Specifically, I will claim that psychological and neuropsychological evidence provides a serious competitor to the broad, ‘single-capacity’ conception of intuition that the views surveyed in section one suggest.

This competitor, which I call the ‘heterogeneous faculties’ approach, hypothesizes diverse psychological underpinnings for different classes of intuitive judgment – thereby suggesting that reliability may need to be assessed separately for moral intuitions, for epistemological intuitions, for logical intuitions, and so forth. A review of empirical evidence supporting this approach will be the focus of section three. The upshot will be that, if the heterogeneous faculties view is indeed correct, the category picked out by ‘intuition’ is hopelessly over-broad. Consequently, the question of the reliability of intuition – and, in turn, the question of the status of intuition-based methodology – appears to have no simple answer. The final section will discuss the consequences of this proposal for current debates on philosophical methodology.

1. Current perspectives on intuition
Though the criticisms and defenses of intuition which have arisen over the past fifteen or so years are quite varied, there is a general – though admittedly not universal – tendency to assume that the intuitions invoked by philosophers\(^{13}\) stand or fall together, and that their doing so will be a consequence of the reliability or lack thereof of some mental capacity called ‘intuition’. This tendency is shared even by philosophers whose views on the nature of this capacity are quite distinct. It will be useful to provide a few quick examples of the phenomenon.

Perhaps the clearest example of the single-capacity view of intuition is found in the work of George Bealer. Bealer claims that intuition forms something like a natural mental kind – that it is a “sui generis, irreducible, natural (i.e., non-Cambridge-like) propositional attitude” (Bealer 1998, 213). Specifically, Bealer holds the view that intuition is an intellectual seeming – a conscious episode in which, upon considering some proposition P, P simply seems to be true. Further, in such seemings, Bealer claims that P presents itself as necessary; it seems as though P must be the case.

In practice, these intellectual seemings are generally identified by reference to their possession of some distinctive sort of phenomenology. In distinguishing intuitions from guesses, for instance, Bealer writes that “guesses are phenomenologically rather more like choices; they are plainly not seemings” (Bealer 1998, 210). A further example – “it does not seem to me that \(25^2 = 625\); this is something that I learned from calculation or a table. Note how this differs, phenomenologically, from what happens when one has an intuition” (Bealer 1998, 210). It appears as though, for Bealer, being an intuition is a

\(^{13}\) The term ‘intuition’, as used in everyday conversation, appears to differ substantially from the usage relevant to philosophy. In particular, the colloquial use of ‘intuition’ covers hunches, guesses, and the like. In this paper, use of the term ‘intuition’ is meant to reflect philosophical use.
matter of having a certain sort of phenomenal character – a mental state is an intuition if and only if it is an intellectual seeming accompanied by a sense of necessity, and such intellectual seemings are distinguished from other states by reference to their phenomenology. A very broad variety of states qualify. Beyond the standard examples provided by reactions to Gettier cases or Twin Earth scenarios, Bealer notes that we have intuitions

“that phenomenal colors are incompatible, that moral and aesthetic facts supervene on the (totality of) physical and psychological facts, that a given determinate (e.g., a particular phenomenal shade) falls under its determinables (e.g., being a phenomenal shade), that the part/whole relation is transitive over the field of regions, or that congruence is a symmetric relation” (Bealer 1998, 211).

Bealer also mentions mathematical intuitions, logical intuitions, and even intuitions about the infinite divisibility of space and time.

Joel Pust offers a similar account of intuition, agreeing with Bealer that intuitions are “a distinct kind of mental state with their own ‘intellectual’ phenomenology” (Pust 2000, 31). Pust’s characterization differs slightly from Bealer’s, however, in that Pust claims that an intuition must only present its content as necessarily true if the intuiter explicitly considers whether that content is necessary. Most importantly for our purposes, however, Pust claims that “intuition is a genuinely basic faculty distinct from sense-perception, and on no worse initial footing than our other natural faculties” (Pust 2000, 119).

Ernest Sosa has offered a somewhat different account of the nature of intuition - one which, nevertheless, appears to espouse something like a single-capacity approach. While Bealer characterizes intuition as a type of mental state wholly distinct from belief, Sosa views intuition as a sort of disposition to have a belief. “On my proposal, to intuit that p is to be attracted to assent simply through entertaining that representational content.
The intuition is rational if and only if it derives from a competence, and the content is explicitly or implicitly modal” (Sosa 2007, 101). For Sosa, then, a state counts as an intuition just in case it is an inclination to belief which does not arise from inference or from perception, but rather from mere understanding of the relevant proposition. As with Bealer, Sosa believes that a great variety of states qualify as intuitions — including our inclinations to believe “that 2+2 = 4; that no sphere is a cube; that nothing is numerically self-diverse” (Sosa 1998, 260).

It is true that Sosa’s characterization of intuition as based on competence or understanding in principle leaves room for such understanding to be generated by several different sorts of processes. In fact, in his (2006), Sosa speculates that “various sources [of intuitive pull] are possible, and the epistemic efficacy of an intuition might depend substantially on its source” (Sosa 2006, 212). Nonetheless, in a later article Sosa discusses intuition in ways that strongly suggest that he has a single, unified capacity in mind.

“Prima facie there is a role for intuition in simple arithmetic and geometry, moreover, but not only there. Just consider how extensively we rely on intuition. Take, for example, any two sufficiently different shapes… you may know perfectly well that they are different. And what you know is not just that the actual tokens are different: you also know that any word token so shaped would be differently shaped from any thus shaped (as you demonstrate the two shapes in turn)… Why deny ourselves a similar intuitive access to the simple facts involved in our hypothetical philosophical examples? That would seem to be the default position, absent some specific objection,” (Sosa 2007, 3).

Throughout that same article, moreover, Sosa refers to the epistemic status of intuition, strongly implying a monolithic view.

While the philosophers mentioned thus far are by and large defenders of the traditional, a priori approach to philosophical methodology, philosophers less sympathetic to the traditional approach have also tended to characterize intuition in a broadly single-capacity vein. These philosophers are less likely to provide explicit
definitions of intuition; indeed, like Sosa, some even entertain the idea that intuitions may have multiple sources. However, they frequently still tend to cast their methodological discussions in terms of intuition generally, suggesting an endorsement of the aforementioned assumption that intuitions, epistemologically speaking, stand or fall together. For example, Robert Cummins has argued that “we should… dismiss philosophical intuition as epistemologically valueless” (Cummins 1998, 125). In a later passage, Cummins writes: “I do not know if something analogous will replace intuition… but something had better replace it. It cannot support any conclusion worth drawing” (Cummins 1998, 126). Brian Weatherson – who espouses a view according to which use of intuition should be restricted, rather than eliminated – also appears to endorse a single-capacity approach. In response to Sosa’s comparison between intuition and perception, Weatherson writes that “there is a distinction to be drawn here, since perception divides into natural kinds, visual perception, tactile perception, etc, and we can use each of these to calibrate the others. It is hard to see how intuitions can be so divided” (Weatherson 2003, 4).\(^{14}\)

Experimental philosophers, who have used empirical studies to question intuition’s philosophical merits, have a somewhat mixed track record with regard to their endorsement of the single-capacity approach. Weinberg, Nichols and Stich (2001) and Machery, Mallon, Nichols and Stich (2004) are fairly careful to restrict their conclusions to epistemic intuitions and to semantic intuitions, respectively; Weinberg et al, for example, claim only that “a sizeable group of epistemological projects…would be seriously undermined if one or more of a cluster of empirical hypotheses about epistemic

\(^{14}\) Janet Levin makes a similar claim, writing that “in most cases, there is cross-modal agreement among the sources of empirical information, in that the deliverances of vision, touch, and audition rarely conflict… philosophical intuitions, on the other hand, comprise a single ‘mode’” (Levin 2004, 197).
intuitions turns out to be true” (Weinberg, Nichols and Stich 2001, 429, emphasis mine). Nonetheless, experimental philosophers also frequently make reference to the evidential status of intuition, full stop. Indeed, Alexander and Weinberg (2007) report the findings of the papers mentioned above in the following manner: “Weinberg, Nichols, and Stich and Machery, Mallon, Nichols, and Stich found that intuitions are sensitive to such factors as cultural, socioeconomic, and educational background” (Alexander and Weinberg 2007, 65, emphasis mine). As a further example of the tendency to treat intuition as a single capacity, Swain, Alexander, and Weinberg write that “We take the growing body of empirical data impugning various intuitions to present a real challenge for philosophers who wish to rely on intuitions as evidence” (Swain et al. 2008, 153). On the other hand, this claim is quickly followed by the suggestion that “philosophers who wish to continue relying on intuitions as evidence begin empirically investigating intuitions about their favorite thought experiments to determine whether, and which, intuitions may be taken as evidence” (Swain et al. 2008, 154, emphasis mine). Further, Weinberg (2007) argues that critics of intuition should avoid targeting intuition as a whole, and instead focus on targeting a particular type of practice, which he dubs “philosophers’ appeals to intuition”.

It does appear that many philosophers, either explicitly or implicitly, take intuition to be a fairly natural, fairly unified type of mental state. Further, though, many philosophers take the evidential status of this type of mental state to be (at least in large

---

15 It is worth noting some prominent exceptions to this approach. David Lewis believed that our intuitions are merely opinions or beliefs (Lewis 1983, x) - a view echoed by Peter Van Inwagen (2001, 149). More recently, Michael Devitt has argued that intuitions are “empirical theory-laden central-processor responses to phenomena, differing from many other such responses only in being fairly immediate and unreflective” (Devitt 2006, 491) – and that while some may be partly innate, many simply reflect accumulated experience. Finally, Timothy Williamson has recently provided a very explicit rejection of the view that intuitions form the primary source of evidence in philosophy, arguing that “what are called ‘intuitions’ in philosophy are just applications of our ordinary capacities for judgment” (Williamson 2004, 109).
part) a matter of reliability. Bealer is quite explicit in claiming that the evidential worth of intuition is to be determined by reference to reliability - “the only adequate explanation [of the evidential status of intuition] is some kind of truth-based, or reliabilist, explanation” (Bealer 1998, 214). Sosa takes reliability to be crucial to an evaluation of evidential worth, as well – “if the appeal to intuition is to help explain in some way how one knows any [of the things one intuits], then intuition must presumably be a reliable ‘source’ of true belief” (Sosa 1998, 262). On the other side of the debate, Weinberg and colleagues write that “intuitions about esoteric cases are in a state of challenge, in that significant evidence against their reliability has been amassed to which no response has yet been found” (Weinberg et al., under review).

I wish to argue that this focus on reliability makes the status of the single-capacity approach to intuition no small matter. Now there is - in general - nothing wrong with using the word ‘intuition’ as a technical term to pick out a group of mental states defined by, e.g., shared phenomenology. What’s more, it may well be that the word ‘intuition’, as used in the standard non-technical sense, in fact refers to just such a category. I am in no sense advocating a form of eliminativism about intuition. However, the mere existence of intuition is not enough. If one wants to maintain a reliabilist approach to the evaluation of evidential sources, then one ought to give serious consideration to whether there exists a plausible alternative carving of mental states which would, if adopted, cast doubt on one’s claims about the epistemological merits of any of the states falling under the term ‘intuition’. In other words, while there may be such a thing as ‘intuition’, one must not simply assume that it is the relevant classification for the epistemological

---

16 Bealer takes reliabilism to apply only to basic sources of evidence, a category which he takes to include perception and intuition.
purposes at hand. Defending this claim, and demonstrating the existence of a plausibly more relevant carving of the mental terrain, will be the goal of the remainder of the chapter.

As a final preparatory note: with the exception of Bealer, the philosophers mentioned in this section are not always fully explicit regarding their commitment to the position I have been characterizing – namely, the position that the philosophical use of ‘intuition’ picks out a unified category of mental states which should be treated as a potential source of evidence, and whose status as an evidential source will depend on its reliability. Though their language has often suggested such an approach, several of the philosophers discussed above would likely be fairly sympathetic to the ‘heterogeneous faculties’ view of intuition I offer in the second half of the paper. If this is right, then so much the better – the aim of this chapter is not to defeat some particular philosopher’s position, but to call attention to an aspect of the intuition debate that I find to be counter-productive.

2. A ‘generality problem’ for intuition?

The following is a well-known problem for reliabilist epistemological theories: in order to evaluate the justificatory status of a belief, the reliabilist must determine whether the process which generated the belief is a reliable one. But plausibly, any given episode of belief formation falls under numerous process types – the cause of my belief that there is a pen on my desk might be said to be an instance of perceptually-based belief formation, of visually-based belief formation, of pen-identification processing, and so on.
How are we to decide which process type is relevant to determining the epistemological status of the belief at hand? This is known as the ‘generality problem’ for reliabilism.  

This type of problem, I think, has serious consequences for every account which takes reliability to be a litmus test for any sort of epistemological category – be it knowledge, justification, evidence, or otherwise. When one uses reliability to determine the epistemological value of any given state, there is a serious possibility that the process type said to produce that state may be characterized too narrowly, or too broadly. If the process type is of the wrong grain, one might fail to accord epistemological value to a state which warrants it. Arguably worse, though, epistemological value may be granted where it ought to be withheld. A worry therefore arises. As noted earlier, philosophers have frequently defended the use of intuition-generated beliefs by appealing to intuition’s reliability. But might intuition be too broadly-characterized, or even too narrowly-characterized? Might we do better to focus on processes of a different degree of generality?

For an introduction to the problem, see e.g. Conee and Feldman (1998) and Alston (1995).

The description I have just given casts intuition as a process. This characterization may seem to be in tension with the characterizations of intuition found in the literature – as can be seen from the previous section, intuition has variously been described as a sui generis propositional attitude (Bealer, Pust), as an inclination to believe (Sosa), or as simply a type of belief (Lewis). On each of these characterizations, it seems most natural to interpret ‘intuition’ as applying to the outcome of a mental process, rather than to a mental process itself. Nevertheless, at least two of the authors just mentioned also explicitly appeal to the ‘reliability of intuition’ – and on most standard versions of reliabilism, reliability is a property of processes. Plausibly, these authors’ usage reflects the fact that the term ‘intuition’ is rather like ‘perception’ – one can speak of perception as a mental process, or one can speak of ‘a perception’ in the sense of a state produced by the process(es) of perception. Alternately, perhaps the claim that intuition is reliable is only meant to indicate that every intuition is produced by a reliable process. However, in the absence of specific claims to the contrary by the authors, the most obvious interpretation of that claim is that every intuition is produced by the same reliable process (type). Finally, perhaps the claim that intuition is reliable should be taken to mean that every intuition possesses the property of being sufficiently truth-tracking (in, e.g., the sense specified by Nozick (1981)). But again, in the absence of specific details from the authors, the most obvious way to defend this claim would be to hypothesize a single process (type) which grants this property to all intuitions it produces. Thus, essentially the same problem remains, regardless of the interpretation given to the ‘reliability of intuition’ – what reason is there to believe that the ‘intuition-generating’ process type is of the appropriate grain?
I want to emphasize that I have no intention of offering a solution to the generality problem. Nor is it my intention to use the generality problem as a premise in an argument against a reliabilist conception of evidence, or even against a reliability-based assessment of the worth of intuition. On the contrary, I take reliabilism to be among the most plausible epistemological approaches currently on offer, and I am quite convinced that reliability should continue to be a central consideration in debates over philosophical methodology. My aim, instead, is merely to call attention to the following strange fact: while there have been many attempts to offer definitions that succeed in capturing most cases philosophers are inclined to classify as intuitions, very little explicit consideration has been given to the question of whether the categories so defined are ones which can appropriately be evaluated for reliability. That is to say, philosophers do not generally attempt to defend their accounts from the accusation that they carve mental states too broadly or too narrowly for reliabilist purposes. This is particularly surprising given the fact that the states we are inclined to classify as ‘intuitions’, when viewed from the standpoint of psychology, appear to be fairly heterogeneous.

---

19 Bealer does briefly mention the generality problem. He notes that the category of mental states including all and only beliefs that there is no largest prime does not qualify as a reliable source of evidence, despite the fact that such beliefs are always true. This category does not qualify, Bealer claims, because only basic sources of evidence are to be evaluated according to a reliabilist criterion; and further, “something can be a candidate basic source only if it is a natural (i.e., non-Cambridge-like) propositional attitude. Intuition, appearance, introspection, belief, desire, guessing, wondering all qualify” (Bealer 1998, 218). I admit to some perplexity over this response. Reliabilism is standardly cashed out in terms of the reliability of mental processes; further, a process reliabilism is suggested by Bealer’s characterization of reliability in terms of the existence of a “strong modal tie between [the candidate source’s] deliverances and the truth” (Bealer 1998, 216). However, Bealer’s aforementioned ‘natural propositional attitudes’ include belief and desire, neither of which are plausibly construed as processes, much less processes of the appropriate grain for reliabilist evaluation. ‘Belief’ makes the list of potential basic sources of evidence – perhaps ‘belief-formation’ is the intended process? But belief-formation is a paradigm case of a process-type which is too broad to be usefully evaluated for reliability. Say it turns out that an individual’s beliefs tend towards 80 percent truth. Does it follow that every last one of her beliefs is justified, by virtue of being produced by the reliable process belief-formation? This seems implausible. After all, this same individual might have (say) a policy of believing everything she reads in tabloid newspapers, and might therefore have only a 20 percent accuracy rate for beliefs regarding celebrity exploits. Surely her tabloid-informed beliefs are not to be counted as reliable.
Even from a standpoint of psychological naïveté, intuitions can seem a motley bunch. There are classificatory intuitions regarding whether x does or does not count as a case of knowledge, a case of pain, or a case of morally permissible action. There are modal intuitions, regarding whether a certain case is possible, or necessary. There are logical intuitions. There are mathematical intuitions. There are intuitions which may not fall under any of these categories, such as the intuition that phenomenal colors are incompatible. Unsurprisingly, there is at least some empirical evidence that these intuitions are not all produced in the same manner.

In the next section, I will review some empirical evidence which I take to strongly suggest a ‘heterogeneous faculties’ view of intuition. If this heterogeneous faculties view turns out to be correct – that is, if it turns out that different types of intuitive judgments are produced by different psychological mechanisms – there may be some reason to think that the question of the reliability of intuition is about as informative, at least for methodological purposes, as the question of the reliability of thinking.

3. Evidence for a ‘heterogeneous faculties’ view

We will now begin a brief tour of some relevant empirical work on types of cognition generally regarded as ‘intuitive’. This section will present two general sorts of considerations which I take to provide good reason to take a ‘heterogeneous faculties’ view of intuition seriously. The first involves brief presentations of the views of researchers who have suggested more or less domain specific models for the types of cognition under consideration. This may seem, *prima facie*, to involve an inappropriate
sort of ‘argument from authority’. However, my aim is not to invoke such models as proof of the heterogeneous faculties approach, but merely to demonstrate that the single-capacity view of intuition cannot be taken as uncontroversial. Insofar as the psychological (and in some cases, philosophical) literature contains serious consideration of the existence of such domain specific processing, it is incumbent on philosophers interested in the epistemology of intuition to at least take such possibilities into account. The primary claim of this paper is simply that philosophers have generally failed to do so, to the detriment of current debates on methodology.

The second type of consideration presented in favor of the heterogeneous faculties approach is a bit more concrete, involving examples of domain-specific cognitive deficits in areas relevant to philosophical theorizing. Such deficits may either be acquired as a result of brain lesions, or they may manifest as a developmental impairment. In either case, the existence of such deficits can provide evidence that certain types of ‘intuitive’ reasoning might be dissociable from other types. If, for example, one’s moral reasoning can become impaired without any effect on one’s logical reasoning, then this provides evidence that moral reasoning is at least to some degree independent from logical reasoning – and may well be subserved by a separate cognitive mechanism. However, it must be mentioned that in such a case the possibility would remain that moral reasoning is dependent on logical reasoning in addition to some other factor – and that,

---

20 There are, admittedly, other possible interpretations. A single ‘intuition mechanism’ might produce both moral and logical judgments, but moral judgments might involve higher demands on the mechanism and might therefore be the first to show impairment when the mechanism has been damaged. Even if this were the correct interpretation, however, there would still be certain epistemological upshots. Consider a parallel case with vision. Though the same basic mechanisms underlie both moderate-distance and long-distance vision, the fact that long-distance vision is more easily impaired (as in myopia) has potential consequences for, e.g., court trials that rely on a witness’s visual experiences as evidence. A court cannot dismiss the possibility of myopia-induced error in a witness’s long-distance identification of a suspect by simply appealing to the on-balance reliability of vision. Mutatis mutandis for the intuition case; any evidence for even single dissociations within the realm of intuition ought therefore to be given serious attention.
therefore, impairment of *logical* reasoning might still result in impairment of *moral* reasoning.

In certain cases, though, further evidence may point to a double dissociation – that is, a case in which function A can become impaired while function B remains intact, *and* (by some other route) function B can become impaired while function A remains intact. Double dissociations provide very strong evidence for independence of function – much stronger than that provided by the single dissociations considered above. But it is important to note that even when a double dissociation is found to be absent, a single dissociation may still provide evidence that two functions should be considered as belonging to separate processes for the purpose of an evaluation of reliability.

Consider the case of reading comprehension and visual acuity. Dyslexia provides evidence that performance on reading comprehension tasks can become substantially impaired without impact upon one’s visual acuity. The reverse does not hold; if one’s visual acuity is sufficiently impaired, one’s performance on reading comprehension tasks will be as well. This performance impairment occurs, of course, because the successful operation of one’s basic visual processes is a *prerequisite* for successful reading comprehension. When the range of circumstances in which one’s vision operates successfully is decreased, the range of circumstances in which one’s reading ability operates successfully is thereby decreased as well. Nonetheless, reading ability and visual acuity are plausibly best considered as separate for epistemological purposes; further, the single dissociation discussed above might plausibly be taken to provide at least some evidence for their separation.
With the preliminaries now in place, let us turn to a presentation of some relevant empirical work. To begin, it is quite plausible that the psychological mechanisms that produce classificatory judgments\footnote{From here on, I will frequently refer to ‘judgments’ rather than ‘intuitions’. I do so because data from the studies being discussed frequently involve verbally reported judgments; though the studies concern areas of cognition philosophers take to involve intuition, the term ‘intuition’ is rarely used. Whether my use of the term ‘judgment’ is warranted may be a matter of debate – philosophers who think intuitions are a type of belief might not object to replacing intuition-talk with judgment-talk, but philosophers who take intuitions to be seemings or inclinations might. Unfortunately, it is not obvious how one would empirically test for the presence of a seeming or an inclination other than by eliciting reports of judgments \textit{caused} by those seemings or inclinations. Thus, I will continue to conduct discussion in terms of judgments; however, this should not be taken as an endorsement of any particular view on whether these judgments are \textit{identical to} or merely \textit{caused by} intuitions.} (judgments about, e.g., whether a given case counts as an instance of knowledge, or of morally right action) are separate from the mechanisms underlying our use of fundamental logical rules or the mechanisms underlying basic mathematical cognition. Neuropsychological and developmental evidence suggests, for example, that there may be one or more domain-specific mechanisms underlying mathematical processing.

McCloskey et al (1985) propose two systems for mathematical cognition – the ‘number-processing system’ and the ‘calculation system’. Each of these systems, in turn, is hypothesized to contain several sub-systems; for example, the number-processing system contains both comprehension and production components. Stanislas Dehaene and his colleagues propose a ‘triple-code’ system, centered on a domain-specific, innate core number system found in the intraparietal sulcus (Dehaene and Cohen (1995), Dehaene (1999), Dehaene et al (2003)). The core system represents quantities in an analog, nonverbal format – rather like a number line - and is useful for approximate calculations and for comparisons of magnitude. However, the brain is also capable of representing numbers in verbal (number word) and visual (Arabic numeral) formats, and these types of representation may be recruited when more precise calculation is required, or when
specific mathematical facts must be stored or recalled. Feigenson et al (2004) suggest that there may be an additional innate system for number comprehension, one which operates on small, precise quantities, allowing for the tracking of small numbers of individual objects – it is this system that allows us to tell, at a glance, the number of items in a group containing between one and four members.

As for evidence from dissociability, the possibility of selective impairment in mathematical ability resulting from brain lesion – a disorder known as ‘acalculia’ – strongly suggests the existence of a more or less dedicated mechanism or mechanisms for mathematical calculation. There is even some evidence for very specific dissociations within the mathematical domain. Dehaene and Cohen (1997), for example, discuss two intriguing patients - one impaired in subtraction, but not in multiplication, and the other in multiplication, but not subtraction. McCloskey et al (1985) report a patient who could accurately compare the magnitudes of quantities presented with number words (e.g. “four”), but not quantities presented in Arabic numerals (e.g. “4”); they also report a patient who could perform basic calculations, but who showed a specific difficulty in retrieving previously learned arithmetical facts (such as 7 x 7 = 49).

At first glance, some of these findings may seem more or less irrelevant to an investigation of the sorts of intuition philosophers are concerned with. After all, the judgment that the group of objects in one’s visual field contains exactly three members (for example) is unlikely to be recruited as evidence for or against any philosophical position; mutatis mutandis for the ability to compare magnitudes presented in Arabic

22 Some cases of acalculia appear to arise due to other deficits, such as language deficits. However, there are other cases that appear to be ‘pure’. Both types of acquired acalculia stand in contrast to the developmental deficit in mathematical reasoning known as dyscalculia. See Ardila and Rosselli (2002) for an extensive overview of these disorders.
numerals. However, other mathematical judgments – such as the judgment that $2+2=4$ – are frequently cited by philosophers as paradigm instances of intuition. At the very least, the judgment that $2+2=4$ appears to fit Bealer’s characterization – it is an intellectual seeming which presents itself as necessary. In fact, I would go so far as to argue that all three of the mathematical judgments mentioned above fit Bealer’s characterization; but even putting this aside, a more important consideration remains.

Consider the following two hypotheses. First, there might exist a faculty of ‘philosophical intuition’ or ‘rational judgment’ which produces epistemological and moral intuitions along with basic arithmetical intuitions such as the judgment that $2+2=4$ - but which is unrelated to the processes which produce, e.g., the ‘non-philosophical’ mathematical judgment that there are three objects in one’s visual field. Alternately, there might exist a mechanism (or, more likely, a cluster of closely related mechanisms) which plays a central role in the production of a variety of mathematical judgments, but which has little to do with (for example) moral cognition. The empirical work discussed above favors the second hypothesis; it thereby casts doubt on the monolithic conception of philosophical intuition.

Let us move to another domain of intuitive thought – logic. Though there is somewhat less empirical work on the psychological underpinnings of the human capacity for logical cognition, at least some researchers have proposed models of deductive reasoning which rely on representations of formal inference rules. Both Braine (1978)
and Rips (1983, 1994) present models which operate via application of stored inference rules similar to those of natural deduction systems in formal logic. More recently, Reverberi et al (2009) have proposed three separate cognitive components for elementary deductive reasoning; one component for rule application, one for tracking the overall structure of the proof, and one for the storing of intermediate conclusions. Reverberi et al also suggest that the processes involved in elementary logical reasoning may differ from those involved in more sophisticated logical reasoning.

Plausibly, systems like the ones the above authors present would be dissociable from systems for other sorts of ‘intuitive’ cognition. Indeed, some evidence for selective impairment in logical cognition exists. Reverberi et al (2009) found that patients with lesions in the medial frontal cortex were impaired in tests of elementary deductive reasoning. They hypothesize “that the overall performance of medial patients could be explained in terms of a deficit in identifying and representing the overall structure of the proof required to solve a deductive problem” (Reverberi et al 2009, 1113). Further, as with mathematics, it appears that selective deficits can arise even within the logical domain. In one particularly fascinating study, Waltz et al. (1999) found that patients with lesions in the prefrontal cortex showed a selective impairment in their ability to perform logical inferences involving reasoning with multiple relations – e.g., in their ability to infer that Bill is taller than Sue from the fact that Jane is taller than Sue and Bill is taller than Jane. Waltz et al take this to be evidence for a specific cognitive ability, central to

---

25 This impairment was not attributable to impairment in verbal comprehension. The deficits of medial frontal patients were also not attributable to deficits in working memory; by contrast, left lateral frontal patients exhibited deficits which appeared to be at least partially linked to deficits in memory. Patients with lesions to the right lateral frontal cortex were unimpaired.

26 Waltz et al note that these deficits were not attributable to deficits in semantic knowledge or episodic memory.
complex logical reasoning, involving the integration of multiple relations – an ability for which the prefrontal cortex may play a central role.

There is at least some evidence, then, that mathematical and logical intuitions are each subserved by at least one – and plausibly multiple – proprietary mechanisms. By contrast, it is rather less plausible that these judgments are produced by some general-purpose ‘intuition’ mechanism shared with classificatory judgments regarding what counts as knowledge, free will, and the like. The ‘heterogeneous faculties’ hypothesis can be taken further, however. There is some preliminary evidence that there may even be separate mechanisms underlying different sorts of classificatory judgments. Though there has not yet been much work on, say, epistemological cognition, there are at least two philosophically relevant areas for which empirical research indicates some degree of domain specificity: moral cognition and folk psychology.

Several researchers have suggested that our moral classifications may be guided by domain-specific mechanisms. Sripada and Stich (2006) suggest that humans possess psychological mechanisms enabling the acquisition and deployment of norms present in their communities. Specifically, Sripada and Stich propose the existence of two innate mental mechanisms devoted to moral cognition. The first, the acquisition mechanism, identifies information regarding norms present in the local environment. The implementation mechanism stores this information, and uses it to generate motivation to both comply with the norms and punish those who violate them. In a somewhat different approach, Hauser (2006) and Mikhail (2007) suggest the existence of an innate, universal ‘moral grammar’ underlying our normative judgments. Like Chomskian universal grammar for language, this moral grammar provides a common base underlying the great
variety observed in culturally local moral systems. And, as with language, this moral grammar places inherent restrictions on the possible range of moral intuitions.

There is a fair amount of evidence that the human capacity for moral judgment can be selectively impaired. The disorder of psychopathy produces subjects with a marked lack of empathy who frequently engage in remorseless amoral behavior. Further, although psychopathy is not associated with any particular deficits on standard cognitive tests, Blair (1995) provides evidence that psychopathic patients have difficulty distinguishing morally wrong actions from actions which merely violate convention. The ability to draw a distinction between moral and conventional violations is often invoked by psychologists as a test of development in moral cognition\textsuperscript{27} - for, unlike conventional violations, moral violations tend to involve harm to others and are typically judged to remain wrong even in the absence of authority-mandated rules prohibiting them. Psychopaths appear to be blind to these differences – for example, the subjects in Blair’s study were much less likely than controls to cite harm to another when asked why a given act was morally wrong.

Even within the realm of moral judgment, evidence suggests that dissociations may occur; and that, therefore, multiple distinct processes may be at work. An fMRI study by Greene et al. (2001) revealed distinct areas of activation when considering ‘impersonal’ moral dilemmas (such as the traditional trolley problem, in which one must choose whether to flip a switch to divert a trolley from a track where it would kill five persons to a track where it would kill one) versus ‘personal’ moral dilemmas (such as the footbridge variant, in which one must choose whether to push a single large man off of a

\textsuperscript{27} See e.g. (Turiel 1983), Smetana(1993). However, see Kelly et al. (2007) for a critique of the relevance of the distinction to moral cognition.
footbridge to stop a trolley from killing five). In a similar vein, Koenigs et al. (2007) found that patients with damage to the ventromedial prefrontal cortex make abnormally utilitarian judgments in response to emotionally charged moral dilemmas; specifically, in dilemmas where aggregate welfare could be increased by performing an emotionally aversive act (such as in the footbridge dilemma), VMPC patients were more likely than controls to say that they would perform the aversive act. By contrast, the VMPC subjects provided normal judgments on other moral dilemmas.

Finally, Cushman (2008) provides evidence that judgments of wrongness and judgments of blameworthiness differ in their reliance on information about mental states. For instance, subjects are much more likely to blame an agent for an irresponsible action when someone is actually harmed, as opposed to a case where an irresponsible action resulted in no negative consequences. This holds even when both cases involve an agent who intentionally acted in an irresponsible manner. Judgments of wrongness, by contrast, rely heavily on information about intent, and largely disregard consequences. Cushman takes this asymmetry to suggest that the two types of judgment are underwritten by distinct cognitive processes, hypothesizing that the intent-driven, mental state-based process emerges later in development than does the simpler consequence-driven process. Cushman also hypothesizes that the two processes “act competitively to determine our [moral] judgments” (Cushman 2008, 355).

In addition to moral cognition, there may also be one or more specific cognitive mechanisms involved in our folk psychological judgments about mental states. There is a vast literature on the mental processes by which we understand the beliefs and desires of others – an ability often referred to as ‘mindreading’. Much of this literature proposes a
distinctive, domain-specific mechanism responsible for mindreading abilities. Baron-
innate mechanism that allows for ‘second-order’ representations – representations, e.g.,
of the mental representations of other people. This mechanism follows a consistent
developmental path among normal children, allowing them to pass tests requiring an
understanding of the possibility of false beliefs at around four years of age.\(^{28}\) Nichols and
Stich (2003) propose a somewhat different model of folk psychological cognition; their
model relies on tacit theory as well as on the ability to simulate the mental processes of
others. Like those of the authors discussed above, the account provided by Nichols and
Stich is more-or-less domain specific. Alvin Goldman (1989, 1992, 2006) presents a
theory of mindreading that relies primarily on simulation, with only a limited role for
theory. Goldman’s is not a purely domain-specific account, since he takes simulation to
underlie other sorts of cognition, as well; however, it still suggests some degree of
heterogeneity in intuition, for Goldman does not take simulation to underlie logical or
mathematical cognition.

The hypothesis that folk psychological cognition involves some degree of domain
specificity is also supported by recent neuropsychological work. Fletcher et al. (1995)
found several specific brain regions which showed increased activation during
presentation of vignettes requiring mindreading compared to non-social control stories.
Even more tellingly, Saxe and Kanwisher (2003) demonstrate increased activation in the
tempo-parietal junction when reading vignettes involving false beliefs, but not when

\(^{28}\) The ‘false belief’ task is a standard test of folk psychological development – see Wimmer and Perner
(1983) for the initial formulation of the task, as well as Wellman et al. (2001) for a meta-analysis of studies
involving the task.
reading vignettes involving false non-mental representations (i.e., a photograph that no longer accurately represents its subject).

There is also evidence that our capacity for mindreading can be selectively impaired. Autistic patients famously have difficulty with social interaction and communication, and children with autism generally fail the standard false-belief task psychologists use as a test for development of mindreading abilities. The mindreading deficit associated with autism is not obviously linked with disproportionately high deficits in other ‘intuitive’ capacities – for example, autistic patients do not seem to have difficulty on the moral/conventional tasks described earlier (Blair 1996). Indeed, autistic patients occasionally even exhibit savantism, which frequently manifests in the form of increased mathematical abilities – a prototypically intuitive kind of cognition. Further, autistic patients may not even exhibit impairment in all mind-related cognition – see e.g. Baron-Cohen (1991) for evidence that autistic individuals understand that emotions can be caused by frustrated desires, but do not understand that they can be caused by false beliefs.

Admittedly, a fair bit of caution is necessary in assimilating philosophical intuitions about beliefs and other mental states to the mindreading capacities just discussed. It may turn out that the cognitive processes underlying our judgments about the actual mental states of others do not underlie our modal judgments about the mind – our judgments about, e.g., whether philosophical zombies are possible. However, there are some cases where a connection between mindreading and philosophical intuitions

29 There has been some interesting recent research into the nature of folk intuitions on consciousness – see for instance Knobe and Prinz (2008), and Sytsma and Machery (2009). However, I am unaware of any findings as of yet on the dissociability (or lack thereof) of consciousness intuitions, nor am I aware of any authors who propose domain-specific models of this type of reasoning.
about the mind seems fairly plausible. It is quite unclear, for instance, whether even a high-functioning autistic subject would share the mainstream philosophical intuition on Frank Jackson’s “Mary” case. That case, after all, involves intuitions about whether Mary would come to have new knowledge – indeed, new beliefs – upon exiting her black-and-white cell. As a further example, it is unclear whether an autistic subject would share the mainstream intuitive reaction to John Searle’s Chinese room; plausibly, a deficit in understanding the beliefs of others would affect the subject’s ability to judge whether the man in the room understands Chinese. More empirical investigation is of course needed; but for current purposes, all that is required is plausibility, not proof. The possibility of one or more separate mechanisms underlying certain intuitions about mental states, at the very least, cannot be discounted out of hand.\footnote{It is worth noting that Jennifer Nagel (2007) discusses the false belief task in relation to the development of epistemological intuition, rather than intuitions about the mind. Indeed, it seems plausible that epistemological cognition, rather than being fully independent from cognition about mental states, might be dependent upon it (though the reverse may not hold).}

I wish to conclude this section with an important qualification. A large proportion of the findings reviewed in this section suggest that the ‘intuition’ category might more helpfully be subdivided along broadly content-based lines – resulting in categories like ‘moral intuition’, ‘epistemological intuition’, and so forth. This is indeed one potential subdivision of the intuition category, but I wish to emphasize that it is not the only possible approach. In fact, the evidence reviewed above suggests that the actual situation may be far more complicated; even within domains like moral or logical cognition, substantially different types of processing may be involved. If this is right, the epistemological features of philosophical methodology may be even more complex. Rather than commit myself to any particular implementation of the heterogeneous
faculties approach, however, I would merely like to emphasize that the psychological and neuroscientific literatures contain a wealth of findings of great relevance for philosophers interested in the epistemology of intuition. Alternate interpretations of these findings can only enhance our understanding of the epistemology of philosophical methodology.

4. Consequences for current debates

Though the data reviewed in the previous section are not of course conclusive, they are suggestive. If there are in fact separate psychological mechanisms underlying different subsets of ‘intuitive’ judgments, then philosophers in the business of arguing for or against intuition need to provide arguments as to why reliability assessment should not target these mechanisms individually, rather than targeting ‘intuitions’ in the broad, single-capacity sense commonly advocated. If the mechanism underlying logical cognition differs from the mechanism underlying moral cognition, then the reliability of the one likely has little bearing on the reliability of the other – a reliabilism targeting intuition generally, however, cannot capture this.

It is true that, in principle, nothing rules out the possibility that each of these individual subtypes will be independently found to be reliable. Nor does anything rule out the possibility that each will be found unreliable. But in order to argue for either thesis, one would have to examine the relevant mechanisms separately, rather than attempting to argue for the reliability of intuition generally. For some styles of argument, this will present little trouble. As an example, consider the recent data on cross-cultural variation in epistemological and semantic intuitions found in Weinberg, Nichols and
Stich (2001) and Machery, Mallon, Nichols and Stich (2004). These data have often been taken to pose a broad, general threat to the use of intuitions in philosophy. However, on the heterogeneous faculties approach, experimental findings on cross-cultural variation in epistemic intuitions (e.g.) will primarily bear weight on the question of the reliability of epistemological intuitions. Such findings will give us little direct evidence as to the reliability of logical intuitions or of moral intuitions. However, the general approach of using empirical findings on variation to challenge philosophical methodology is not thereby undermined; if an experimentalist suspects unreliability in other domains, she may simply run more studies.

On the other hand, some styles of argument will be less amenable to reformulation in a heterogeneous faculties framework. In particular, ‘self-defeat’ arguments, which aim to show that a critic of intuition must tacitly rely on intuition (and that intuitions must therefore be evidence), will become problematic. Bealer (1993), for example, argues that in order to formulate any successful theory of justification, one must invoke certain basic epistemic classifications, or ‘starting points’. These starting point classifications include judgments regarding what does and what does not count as an experience, an observation, an explanation, a valid argument, and so on. Any philosopher intending to reject intuition as an evidential source, Bealer argues, must provide a theory of justification that would license such a radical departure from our standard epistemological procedures. Providing such a theory of justification is impossible without invoking the starting points just mentioned; but such starting point judgments, Bealer claims, are invariably arrived at via intuition. It is therefore impossible
to defend the rejection of intuition without appeal to intuition – and, it is argued, intuitions must therefore be evidence.

On a heterogeneous faculties approach, however, this argument does not prove that intuitions must be evidence. If there are – as I claim – multiple processes underlying the production of intuitive judgment, such an argument at best shows that epistemological intuitions must be evidence. In fact, if epistemological cognition turns out to be underwritten by multiple processes, it may prove less – it may prove merely that classificatory intuitions about very basic epistemological categories must be counted as evidence. In either case, it cannot be used to defend the evidential status of intuitions in, e.g., the moral domain, nor the evidential status of ‘intuition’, full stop.

Another argument of this same ‘self-defeat’ type has been proposed by Joel Pust (2000). Pust’s target is what he calls ‘explanationist skepticism’ about intuitions. An example of such ‘skepticism’ is provided by an argument found in Harman (1977), which runs broadly as follows. Harman claims that the best explanation of our having the moral intuitions we do does not advert to the truth of the intuited moral propositions – instead, it adverts to various psychological facts. Because of this, we should not take the occurrence of moral intuitions as evidence for the truth of any moral facts corresponding to those intuitions. Pust argues that this argument relies on an ‘explanationist’ criterion of justification as a premise, which runs something like this: aside from propositions describing the occurrence of one’s observations and intuitive judgments, one is justified in believing only those propositions which form part of the best explanation of the occurrence of one’s observations and intuitive judgments. Because of its reliance on this criterion, Pust claims, Harman’s argument undermines itself.
The possibility for self-defeat, Pust claims, arises from the fact that the explanationist criterion of justification can itself only be justified by appeal to intuition. However, Pust further claims that intuition quite generally falls afoul of the explanationist criterion. Thus, the explanationist criterion is deemed unjustified by its own lights. Arguments like Harman’s fail, according to Pust, because they “treat intuiteds as evidence for a principle allowing only intuitings to count as evidence” (Pust 2000, 87). Further, the explanationist argument – indeed, any argument – relies on an ability to grasp an argument’s validity – an ability grounded in intuition. “Such an appeal to intuition, again, however, violates the explanationist principle,” (Pust 2000, 87).

The assumption seems to be that an explanationist like Harman must claim that all intuitive judgments are such that their contents are not invoked in the best explanation of their occurrence; however, Pust provides little argument as to why this should be so. A heterogeneous faculties approach would, in fact, give Harman plenty of room to maneuver here – he could quite easily claim that the sorts of psychological facts that form the best explanation of the occurrence of our moral intuitions have little to do with the best explanation of the occurrence of our epistemological intuitions. On the heterogeneous faculties approach, a debunking argument against moral intuitions no more affects the assumption that epistemological intuitions are generally best explained by the truth of their contents than it does the assumption that visual observations are generally best explained by the truth of their contents.31

---

31 This should of course not be read as a claim that epistemological intuitions are, in fact, in better epistemological shape than moral intuitions – that remains open to investigation. It is merely an argument that one cannot assume that the various types of intuition have the same epistemological features; one cannot take Harman’s claims to undermine intuition simpliciter, any more than one can take Pust’s arguments to support intuition simpliciter.
Whether Harman’s argument is truly self-undermining depends on whether the best explanation of the occurrence of our epistemological intuitions fails to advert to their truth. Harman has claimed that the best explanation of our moral intuitions need only advert to certain psychological facts; however, he has not claimed that the same holds for epistemological intuitions. If the best explanation of our epistemological intuitions does advert to their truth, then there seems to be no trouble in applying the ‘explanationist’ criterion of evidence across the board – epistemological intuitions will pass the test, and thus the justification of the explanationist criterion itself by intuition will be not be self-undermining.

To be fair, the two self-defeat arguments just discussed do function perfectly well against what we might call ‘radical’ anti-intuitionism – the view that no intuitions are admissible as evidence. And it may be that a repudiation of radical anti-intuitionism was the authors’ only intent. But in many cases, arguments such as these have been deployed against philosophers like the experimentalists discussed above, whose arguments against intuition continue to bear weight on methodological issues once they are divorced from any radical tendencies. The Weinberg, Nichols, and Stich findings on cross-cultural variation on Gettier cases, for instance, present a serious challenge to philosophical method even once a heterogeneous faculties view of intuition has been adopted – for they suggest that Gettier intuitions are sensitive to a feature (cultural

---

32 However, as mentioned above, Bealer does appear to explicitly conclude from his self-defeat argument that intuitions must be evidence. In his (1998) he writes that “in ‘The Incoherence of Empiricism’ I argued that [the rejection of intuition leads] one to epistemic self-defeat. In this chapter, I will just assume that these arguments succeed and that we cannot coherently deny that intuitions have evidential weight” (p. 214). In case one is tempted to interpret this as a mere rejection of radical anti-intuitionism, note that the very next sentence reads: “What explains why intuitions are evidence?” (p.214). Pust is less bold than Bealer in this regard, but does mention in a footnote that he is “inclined to see [the self-defeat argument] as providing one reason to trust intuition more than sense perception, namely, that without it, no coherent epistemology can be constructed” (Pust 2000, 110).
background) which does not obviously affect the truth values of knowledge claims. It is not sufficient to dismiss such arguments by merely discrediting radical anti-intuitionism.

One final point bears mentioning. I have suggested that the question of the reliability of intuition is relatively uninteresting; instead, we should focus on the reliability of much narrower types of cognition. However, one might argue that there is still plenty of sense to be made of the project of determining the overall reliability of intuition. After all, there is plenty of literature devoted to the ‘rationality debate’ – the question of the degree to which humans can be said to be rational in face of certain well-known, common errors on basic reasoning problems such as the Wason selection task. No one takes participants in that debate to be blithely ignoring distinctions between various types of reasoning; they are merely concerned with a higher-level question.33

I agree with the basic thrust of this objection; it is, in one sense, quite reasonable to inquire as to the overall reliability of types of reasoning we generally categorize as intuitive. One might simply be intrinsically interested in the reliability of intuition. But this is not the relevant question for current debates over current philosophical methodology. As an analogy, one might quite reasonably inquire into the overall reliability of perception, as a matter of intellectual curiosity; but when questions arise as to, e.g., the viability of various observational methods in science, it is time to draw our focus much more closely. Similarly, I claim that a narrower scope is appropriate in debates over traditional methodology.

To summarize, I have attempted to argue that the current dominant approach to philosophical methodology, which takes the evaluation of a priori methods to consist in an evaluation of the reliability of a single faculty known as ‘intuition’, is problematic. At

33 This objection was suggested to me by Alvin Goldman.
the very least, a substantive argument must be offered as to why an assessment of reliability should be aimed at intuition generally rather than at several more narrowly-defined capacities. If the heterogeneous faculties approach to intuition bears fruit, however, then is quite possible that the single-capacity view of intuition will need to be abandoned entirely. If this is the case, the very question of the evidential status of intuition may turn out to be irrelevant to debates over the status of traditional philosophical methodology.
Chapter 3

Intuition, Evidence, and the Threat of Skepticism

There is a popular view, occurring implicitly and sometimes explicitly throughout the intuition literature, according to which critics of intuition are in fact purveyors of skepticism. One clear example of this view occurs in Pust (2000), which argues against what Pust terms ‘explanationist skepticism about intuitions’, claiming that such arguments are inevitably self-undermining. Bealer (1992) offers a similar ‘self-defeat’ argument. Somewhat less explicitly, questioning the evidential status of intuition is frequently compared to questioning the status of perception – see for example Bealer (1992) and Sosa (1998, 2006). Most recently, Timothy Williamson pursues this approach in The Philosophy of Philosophy. Williamson argues that philosophers who question the use of intuition are ‘judgment skeptics’, employing arguments which essentially parallel arguments for traditional skepticism about perception. Since we do not take traditional skepticism seriously during most of our philosophical theorizing, we should feel no threat from the judgment skeptic; we are therefore free, Williamson claims, to take the contents of our intuitive judgments as evidence for philosophical hypotheses.

This chapter will focus on the Williamsonian version of the charge of skepticism against critics of intuition. The immediate aim will be to show that it is unclear who, if anyone, fits the mold of Williamson’s ‘judgment skeptic’. My hope, however, is that the morals drawn will be more general; one such moral being that certain arguments against

---

34 Though Williamson regards the word ‘intuition’ as fairly unhelpful, in that it fails to carve out a distinctive category of mental states, I will continue to use the word occasionally for sake of exposition. On my intended usage, ‘intuition’ simply picks out those judgments that critics of intuition take themselves to be criticizing. This primarily includes immediate, first-up reactions to thought experiments. No distinctive cognitive status for such judgments is assumed.
the use of intuition in philosophy are often taken to have far more wide-reaching implications than they actually possess. In particular, the empirically-based arguments of Stephen Stich, Jonathan Weinberg, and other members of the ‘experimental philosophy’ movement - though many might take them to be prime examples of the sort of skepticism Williamson targets - do not plausibly mirror standard arguments for classical skepticism, and do not entail any sort of skeptical consequences. The experimentalists’ arguments, unlike skeptical arguments, do not jeopardize the entirety of our capacity for inquiry – they do, however, raise important concerns about the extent to which intuition should be employed in philosophical theory-building.

1. Evidence in philosophical theorizing

As a preliminary, it is important to be clear on the conception of evidence that Williamson takes to be appropriate to philosophy. There is a loose sense of evidence upon which physical objects like footprints or documents can constitute evidence - this is not the notion of evidence currently under consideration. For Williamson, and for the purposes of this chapter, one’s evidence consists of true propositions to which one has some appropriate sort of epistemic access. The nature of this epistemic access is, of course, a matter open to debate.

Because we can be uncertain as to whether a given proposition is true, or as to whether we stand in the appropriate epistemic relation to it, we are not always in a position to identify our evidence. When two parties dispute the evidential status of a given proposition, the role of that proposition in debate is compromised. If we are
feeling optimistic, however, we might suppose that such disputes can at least in principle always be resolved; we might suppose that something can count as evidence only if all relevant parties are able to agree that the proposition is in fact evidence. Williamson terms this idea ‘Evidence Neutrality’.

The trouble with this idea is that there are a great many propositions – ones which philosophers would like to use as evidence - which do not appear to satisfy Evidence Neutrality. In particular, Williamson claims, a group of philosophers whom he terms ‘judgment skeptics’ refuse to agree that statements involving certain common-sense, intuitive judgments are admissible as evidence. To take Williamson’s example, a widely-accepted proposition like the Gettier case is not a case of knowledge may be rejected by the judgment skeptic as mere cultural bias. What must one do when faced with such an interlocutor? Evidence Neutrality bars you from taking the Gettier proposition to be evidence; you must retreat to propositions both parties accept. One plausible candidate is the proposition expressed by the sentence ‘I intuit that the Gettier case is not a case of knowledge’. Both the judgment skeptic and the non-skeptic can agree on this - the mental states of the debate’s participants are not currently under suspicion.

However, Williamson warns that to retreat in this manner is to risk making the questionable move of ‘psychologizing the evidence’. When one psychologizes the evidence, one commits oneself to the claim that our evidence consists only of our intuitions or other psychological states, and not of their content. Thus, propositions about knowledge and other philosophically interesting phenomena are barred from the evidential base, to be replaced by propositions about our beliefs about knowledge, our intuitions about knowledge, or our concept of knowledge. It is “as though our epistemic
access were only to those belief states and not to the states of the world they are about” (Williamson 2007, 5).

Applying this sort of psychologizing move in response to disagreement about the Gettier proposition puts one in the unhappy position of attempting to argue from a psychological premise (“I intuit…”) to an epistemological conclusion about the nature of knowledge. The gap is not easily bridged. One may be tempted to psychologize the subject matter of philosophy in order to address the problem, arguing that the aim of philosophy is to characterize psychological phenomena like concepts. This, Williamson claims, would be a mistake – philosophy is not primarily concerned with concepts, but with phenomena in the world. We should therefore resist this psychologizing of the evidence; further, we should reject the principle of Evidence Neutrality that led us to it. “Having good evidence for a belief does not require being able to persuade all comers, however strange their views” (Williamson 2007, 212). We should continue to employ the Gettier proposition as evidence, despite the judgment skeptic’s objections.

The strategy just outlined leaves us in a bit of a bind. There have, in recent years, been numerous empirical challenges to the use of intuition-based premises in philosophical argumentation. One prominent example argues that cultural background and socioeconomic status influence responses to thought experiments (Weinberg, Nichols and Stich 2001); another claims that such responses are susceptible to order effects and other cognitive errors (Swain, Alexander and Weinberg 2008). The common strategy of such arguments is to allege that intuitions vary as a function of factors which are irrelevant to the hypotheses under examination. Cultural background of the attributor, for instance, should be irrelevant to the question of whether or not an agent in a Gettier
scenario has knowledge.\textsuperscript{35} Order of presentation of vignettes should be similarly irrelevant to questions about whether the agent described by said vignettes has knowledge. Given that these factors do not affect the truth of knowledge ascriptions, at least some of the observed judgments must be false. I call arguments that take this general form ‘variation arguments’.

Variation arguments are difficult to dismiss out of hand; they at least \textit{prima facie} present us with plausible reasons to doubt the evidential status of intuited propositions. However, as Williamson notes, if these challenges in fact lead us to replace all intuition-based assertions – i.e., assertions about the contents of intuitive judgments - with psychological assertions about the occurrence of certain mental states, we may have a difficult time making philosophical headway. At best, we may be able to come to conclusions about the nature of our concepts. Does this potential obstacle to philosophical progress justify a refusal to countenance the ‘strange views’ of the experimentalists?

There are, of course, philosophers who \textit{do} take their aim to be the characterization of concepts (in the psychological sense); Alvin Goldman and Joshua Knobe, for instance, are quite explicit about doing so (See Goldman 2007, Knobe 2007). Nonetheless, Williamson is correct in his claim that many philosophers are interested not in the concept of knowledge, but in knowledge itself. He is also correct in his claim that arguments from psychological premises to conclusions about phenomena like knowledge are not straightforward. It is, however, worth reminding ourselves that discussion of

\textsuperscript{35} Unless of course one embraces some form of relativism about matters epistemic. This has not, however, been a particularly popular strategy.
intuitions or of other psychological states can be quite relevant when addressing philosophers’ claims about non-psychological phenomena.

Framing debate in terms of statements about mental states does not, in general, imply a refusal to accept any non-psychological statements into the evidential base; instead, it typically reflects a concern with the sources of statements being held up as evidence. The rejection of Evidence Neutrality does not justify wholesale refusal to accept psychologically-inspired critiques of the source of one’s claims. If examination of the source of belief in a particular assertion suggests that some evidence-constitutive epistemic relation does not hold between the asserter and the proposition asserted, this is grounds for a claim that the statement must be barred from debate. One may find grounds to deny that any judgment formed on the basis of the source in question would constitute evidence; alternately, one may find that the source in question fails to provide evidence only in certain unfavorable contexts.

These sorts of maneuvers are commonplace; they occur, for instance, when we doubt a murder witness’s ability to recognize a face in a dark room, on the grounds that perception is unreliable in low lighting conditions. They occur when we reject the use of ‘clairvoyance’, or when we argue that non-blinded experimental studies risk skewed results. Such moves do not automatically ‘psychologize’ the argument. Though they concern psychological states, they are relevant to the evaluation of non-psychological claims. If critics of intuition are guilty of psychologizing the evidence, it cannot simply be in virtue of criticism of this sort; there must be some principle that separates cases like the above, which we might call ‘mundane epistemic criticism’, from the ‘strange views’ we are confident in ignoring.
Williamson would, of course, agree with the claim that psychological propositions are frequently relevant for the evaluation of non-psychological claims. In his discussion of the problem of vagueness, for instance, he notes that although a question such as ‘Was Mars always either dry or not dry?’ is not in any sense about thought or about language, “to answer it adequately one must assess rival theories of vagueness in thought and language” (Williamson 2007, 41). Nonetheless, Williamson seems committed to the idea that critics of intuition do not argue in the manner of the philosopher who concerns himself with language in order to address non-linguistic issues of vagueness. Nor do they merely offer mundane epistemic criticism; according to Williamson, their attacks are more pernicious.

The purported difference lies in the extent to which the rejection or restriction of the targeted source depletes our evidential assets. The judgment skeptic, Williamson argues, resembles a skeptic about perception. In contrast to cases of mundane epistemic criticism, the ‘psychologizing’ of a perceptual skeptic is genuinely incapacitating. A skeptic about perception refuses to admit into debate any statements about the contents of our perceptions, on the grounds that we cannot be certain that our perceptual capacities provide an appropriate link to truth in any circumstances. Such broad skepticism provides too little ground to stand on – progress is impossible. “One regards [the skeptic’s] restricted evidence base as too wilfully impoverished to constitute a reasonable starting-point for evaluating the propositions in which one is interested” (Williamson 2004, 124). It is thus plausible that wholesale perceptual skepticism, unlike mundane epistemic criticism, is subject to dismissal by virtue of the failure of Evidence Neutrality.
Williamson claims that both the perception skeptic and the judgment skeptic are guilty of psychologizing the evidence in a manner that puts unreasonable restrictions on permissible evidence.

“Skepticism about perception typically narrows one’s evidential base to one’s present internal mental state… Judgment skepticism narrows and internalizes our evidential base in a similar way without going as far as skepticism about perception, since typically it treats other people on a par with oneself, and other times on a par with the present” (Williamson 2007, 226).

The judgment skeptic, Williamson says, “insists that I have as evidence at most the fact that it non-perceptually appears to me and others that the subject in a Gettier case lacks knowledge” (Williamson 2007, 226). If this is in fact the sort of argument the critic of intuition offers – if she is, in fact, a judgment skeptic – then she is plausibly subject to the same dismissal as the skeptic about perception.

Thus far, we have determined the following. If the critic of intuition does indeed resemble the skeptic about perception with regard to the extent of evidence rejected, the falsity of Evidence Neutrality allows us to refuse to give any credence to the ‘skeptical scenarios’ the critic presents. If, however, the critic of intuition offers merely mundane epistemic criticism, then the rejection of Evidence Neutrality does not warrant continued use of the targeted propositions in philosophical argumentation; at least, not without some further argument defending the source of one’s belief in said propositions.

Critics of intuition have not generally characterized themselves as endorsing as severe a restriction on evidence as the skeptic about perception; they oppose the current widespread use of intuition in arguing for philosophical claims, but they do not call for the banning of all propositions which do not concern facts about our mental states. Williamson, however, claims that their arguments generalize – the rejection of intuition
leads to the rejection of non-psychological judgment across the board. I will consider this claim in the next section.

2. Intuition and concept application

Though his targets might identify themselves as skeptics about the use of intuition, Williamson notes that that there is no widely-accepted account of the nature of states falling under the term ‘intuition’. ‘Intuitive’ judgments, in fact, do not appear to enjoy any particular distinguishing feature – unlike processes like perception and memory, intuition does not form a distinct mental kind. Intuitive judgments tend to employ the same cognitive capacities as other, everyday judgments.

As an example, consider the Gettier thought experiment. In evaluating the Gettier scenario, Williamson claims, we do not brutely exercise some ‘faculty of intuition’ - we in fact perform a number of cognitive tasks. First, we determine that the case is possible. Second, we evaluate a counterfactual conditional of the following form: *if the Gettier scenario had occurred, then the subject would have had justified true belief without knowledge*. Third, we judge that the counterfactual conditional forms a counterexample to the claim *knowledge is justified true belief*. Williamson pays particular attention to the second of these tasks, and emphasizes the ubiquity and reliability of such counterfactual evaluations in everyday cognition. As an example of this general ability to evaluate counterfactuals, Williamson offers a case in which a subject watches a rock slide down a mountainside and get caught in a bush. The subject is quite capable of accurately forming a counterfactual judgment of the following form – ‘If the bush had not been
there, the rock would have landed in the lake’. On Williamson’s account of the
psychology of counterfactuals, the subject does this by imagining or simulating the
antecedent, and then employing, ‘offline’, the same capacities that are used when he
predicts the trajectory of a fast-moving object currently passing by – i.e., he employs
something like a hard-wired folk physics. In general, evaluating counterfactual
conditionals involves offline usage of the same cognitive capacities used in evaluation of
actual cases.

Of course, the standard critic of intuition does not tend to focus her objections on
the modal nature of the judgment in the Gettier case; instead, she questions the subject’s
ability to accurately classify the case as one of knowledge. The reliability of our folk
physical predictions and other similar predictive faculties does not directly affect this
challenge. Indeed, Williamson notes that in answering a question like ‘If twelve people
had come to the party, would it have been a large party?’ one doesn’t “imagine a party of
twelve people and then predict what would happen next” (Williamson 2007, 151).
Classificatory judgments of this sort involve a different kind of cognitive task than
physical predictions, though one that may still be run ‘offline’ in the evaluation of
counterfactuals. Nevertheless, Williamson still claims that classification in the Gettier
case doesn’t involve any peculiarly philosophical intellectual ability - it involves “the
same capacity to classify empirically encountered cases with respect to knowledge as we
use when, for example, we classify a politician as not knowing the truth of his claims
about terrorists” (Williamson 2004, 112). Thus, ‘intuited’ responses to thought
experiments rely on general-purpose cognitive capacities – including the capacity to
make classificatory judgments.
Given this view on intuition, Williamson proposes the following portrayal of the critic of intuition:

“Skepticism about intuition consists not in skepticism about a special kind of judgment but in a special kind of skepticism about any judgment. That skepticism does not target the distinctive features of perception, memory, testimony, or inference. Rather, it targets our practices of applying concepts in judgment” (Williamson 2007, 220).

A generalized skepticism about concept application would certainly threaten the sort of intellectual paralysis that Williamson is concerned with; for, plausibly, every judgment involves some use of concept application. Indeed, a thoroughgoing skeptic about concept application could not even take refuge in psychological statements like ‘I intuit that the Gettier case is not a case of knowledge’, or ‘It appears as though the cat is on the mat’, for her skepticism would apply even to her ability to classify those states as intuitings and appearings.

Williamson proceeds to claim that judgment skeptics rarely present themselves as skeptical of all instances of concept application; instead, they tend to focus on concept application within a particular context.

“In a context that concerns folk psychological ascriptions of belief and desire, Paul Churchland and other eliminativists about such mental states are judgment skeptics. In a context that concerns ordinary geographical judgments, Terry Horgan and other eliminativists about mountains are judgment skeptics. Such skeptics question our standards for applying ordinary concepts both in experience and in thought” (Williamson 2007, 220).

‘Judgment skeptics’ such as Churchland and Horgan aim to construct ‘skeptical scenarios’ in which the ability to apply a given folk concept confers an evolutionary advantage, yet nonetheless fails to reflect the underlying reality. Without naming names, Williamson adds that other judgment skeptics may offer more sociological ‘skeptical scenarios’ in which culturally-variable concepts like ‘knowledge’ serve important social functions despite their general inadequacy. Scenarios such as these, Williamson claims, aim to “make massive error a genuine possibility” (Williamson 2007, 221). The claim
being attributed to the judgment skeptic, then, is that we simply manage to get by despite pervasive falsehoods - at least until the time comes to settle down and do our science.

It is at this point, however, that Williamson’s target becomes unclear. Two distinct groups appear to stand accused of judgment skepticism. The first group consists of philosophers who take an eliminativist stance about some folk category or another, but do not necessarily argue against intuition more generally. The second group consists of experimentalists, who invoke experimental data about e.g. cultural variation to make more general condemnations of traditional philosophical methodology. However, even experimentalists have not taken their conclusions to impugn concept application across the board. It is not yet particularly obvious how either of these sorts of context-specific arguments might go beyond mundane epistemic criticism, unavoidably generalizing to generate radical skepticism.

There appear to be two primary argumentative features which Williamson takes to be jointly sufficient to enable a generalization to overall skepticism. The first has already been mentioned – the offering of a possible scenario in which our current concept-application practices (for one or more folk concepts) result in massive error. The second argumentative feature is the demand that our evidence be taken to be neutral between the truth of the ‘skeptical scenario’ and the truth of the commonsense worldview.

Williamson accuses judgment skeptics of this second strategy at several points:

“[Judgment skeptics] try to put defenders of a piece of common sense into the position of arguing for it over the judgment skeptical scenario from a starting point neutral between the two alternatives, just as skeptics about the external world do” (Williamson 2007, 223).

And further:

“A judgment skeptic argues that our evidence is neutral between the ordinary hypothesis that there are mountains and the skeptical hypothesis that there are no mountains, but instead only complex microphysical events” (Williamson 2007, 224-225).
As part of this latter move, Williamson imagines the judgment skeptic psychologizing the evidence - insisting that the evidential starting point include only propositions about such psychological states as, e.g. being inclined to judge that there are mountains. The judgment skeptic then questions our right to take our inclination to judge that there are mountains and treat it as good evidence that there are, in fact, mountains.36

That the combination of these strategies might lead to generalized skepticism is fairly clear. Imagine a philosopher puts forth a possible scenario which would entail massive error in a particular domain D; i.e., if the scenario were true it would entail that our judgments regarding D are overwhelmingly false. Further, the philosopher claims that our evidence is neutral between the proposed scenario and a commonsense worldview. All the evidence we really have about D, the philosopher argues, is the fact that we tend to make judgments whose truth would support the commonsense view of D; but this is consistent with the skeptical scenario, as well. When presented with such an argument, the following question naturally arises: why should we think we are radically deceived regarding D, but not similarly deceived in other contexts? If the argument is simply that our psychological evidence is consistent with both a commonsense

---

36 This mode of argument, Williamson claims, relies on an illicit “appearance principle”. According to this principle, one should be confident in some proposition P on the basis of an appearance that P (where this is some non-perceptual sense of ‘appearance’ meant to capture most cases covered by the term ‘intuition’) only if its appearing that P is good evidence that P. Williamson devotes a fair bit of time to demonstrating that the appearance principle generates an argument for widespread skepticism, and concludes that judgment skeptics should not rely on such principles in their reasoning. The details of this argument are not relevant for our purposes, however. No member of either group mentioned above has, to my knowledge, explicitly accepted or endorsed such a principle. The principle has some prima facie appeal as a justification for a rejection of intuition; but then, it has some prima facie appeal as a justification for a rejection of any problematic source of evidence. One can imagine parallel principles employed in rejections of clairvoyance, or of perceptions in dark rooms. Surely, a principle by which to evaluate questionable sources of evidence would be of great value for the critic of intuition, as well as for those offering uncontroversially mundane epistemic criticism. But we don’t require any such principle be offered before any reasonable evidential criticism is taken seriously.
worldview as well as the imagined skeptical scenario, then doesn’t a parallel argument apply in every domain?

One can, it appears, generalize to overall skepticism given both a suggestion of massive error in some domain combined with a claim that the relevant evidence – which consists solely of psychological facts – is neutral between the error scenario and the commonsense worldview. I am of the opinion, however, that neither of the groups Williamson mentions employs both of these argumentative features. If this is right, it is unclear that there are any judgment skeptics. The main aim of this chapter is to defend the variation arguments offered by experimentalists; I will, in the next section, begin to make the case that they do not fit the judgment skeptic mold. However, it is worth briefly discussing the accuracy of this portrayal when applied to the eliminativist, as well.

To start, we must admit that eliminativists deal in scenarios suggesting massive error. Eliminative positions are ontological – they question the existence of a particular type of entity, property, etc. Such views entail that our positive classificatory judgments involving the target category are in error. Any proposition of the form ‘John believes that…’ is false if eliminativism about propositional attitudes is true, because the non-existence of beliefs entails that no one believes anything. Any proposition of the form ‘x is a mountain’ is false if eliminativism about mountains is true. An eliminativist will, therefore, refuse to accept such propositions as evidence – she will at best accept as

---

37 By ‘positive classificatory judgments’ I simply intend to exclude negative judgments of the form “John does not believe…”, which may still be true under eliminativism.

38 Some versions of eliminativism might consider these propositions meaningless and/or truth-valueless rather than false. Regardless, the propositions will certainly not be true – therefore, they will not be candidates for evidence.
evidence propositions regarding mountain-appearances\textsuperscript{39}, while claiming that all such appearances fail to reflect truth.

The scenarios suggested by eliminativists entail massive error in a particular domain; they entail that our judgments within that domain are overwhelmingly false. This purported failure in our categorization practices might even seem to threaten a wide-ranging skepticism about concept application, especially if it were supported merely by a claim that such scenarios are consistent with the psychological evidence. Fortunately for the eliminativist, Williamson’s second criterion for judgment skepticism does not properly characterize eliminativist arguments. Eliminativists simply do not claim that our evidence is \textit{neutral} between their ‘skeptical scenarios’ and the commonsense worldview; rather, they claim our evidence tells against the commonsense view.

These philosophers do not, like the traditional skeptic, merely suggest the possibility of a skeptical scenario; they provide evidence that their scenarios are actual. Churchland (1981), for instance, argues that folk psychology fails to adequately explain certain central mental phenomena like mental illness, sleep, and memory; that it has failed to advance significantly in several thousand years; and that the applicability of folk psychology has in fact narrowed – in ‘primitive’ cultures, mental states were routinely applied to elements of nature like rivers. One may fail to be convinced that such considerations show that folk psychology is a false theory, but there is no doubt that Churchland is arguing not for possibility, but for actuality. Skeptics about perception, by contrast, have failed to provide so much as a single evil demon footprint.

Williamson mentions that the judgment skeptic differs from the perceptual skeptic in this regard, but doesn’t seem to view this as an important difference. However, it is

\textsuperscript{39} An eliminativist about propositional attitudes might hesitate to accept even appearance claims.
not clear why he thinks that the eliminativist’s argument – that there are reasons to believe that we are in a scenario that discredits our ability to correctly apply concepts in a particular domain – entails commitment to the demand that we withhold all concept application on grounds of the possibility of such scenarios in other domains. An argument to the effect that we are in a ‘skeptical scenario’ that prevents us from making accurate visual judgments in certain contexts (e.g. dark rooms) does not commit one to the argument that all perceptual evidence is illegitimate until proven otherwise. Eliminativist arguments appear to be cases of merely mundane epistemic criticism, rather than springboards to skepticism.

3. Variation arguments and massive error

What about the experimentalists? Though Williamson appears to lump experimentalist arguments into the same broad category as eliminativist arguments, there are important differences between the two. The eliminativist argues, on empirical grounds, that we are operating with a radically defective concept or concepts – that entities corresponding to these concepts do not exist - and that we are therefore in massive error with regard to concept application in a particular domain. The experimentalist argues, on empirical grounds, that intuitive judgments in response to thought experiments exhibit some epistemologically undesirable feature such as sensitivity to cultural background. The presence of this feature is then taken to present a challenge to current philosophical practice.
With regard to Williamson’s second criterion for judgment skepticism, the experimentalists bear some resemblance to the eliminativists; for neither group claims that our evidence is neutral between the scenarios they propose and the commonsense worldview. Like eliminativists, experimentalists offer positive evidence for error with regard to particular concepts – the cross-cultural findings on Gettier intuitions offer such evidence for the concept of knowledge. However, there is one important difference; while eliminativists are often content to critique only certain selected concepts, the broader ambition of many experimentalists is to use their empirical findings to suggest a more general inadequacy in intuition as a whole, one which potentially undermines current philosophical method. Given Williamson’s fairly convincing claim that there is no principled way to distinguish between intuition and other general-purpose intellectual capacities like concept application, experimentalists may in fact seem at risk of impugning concept application across the board.

Indeed, if a philosopher of the experimentalist stripe were an eliminativist as well, an accusation of generalized skepticism might be quite apt. Such a philosopher might claim that his findings implied that no entities corresponding to philosophically important concepts exist – that we are in massive error wherever intuition is applied. If Williamson is right, this would suggest that we are in massive error whenever we apply concepts – leaving us with a sorely insufficient supply of genuine evidence. In practice, however, these two categories of philosophical trouble-making do not typically overlap. Unlike the

\[40\] Weinberg et al, for example, write “it is difficult to see why a process that relies heavily on epistemic intuitions that are local to one’s own cultural and socioeconomic group would lead to genuinely normative conclusions” (Weinberg et al. 2001, 36); Swain, Alexander, and Weinberg write that “[the observed] instability undermines the supposed evidential status of these intuitions, such that philosophers who deal in intuitions can no longer rest comfortably in their armchairs” (Swain et al. 2008, 138).
eliminativists, experimentalists do not meet the first of Williamson’s criteria for judgment skepticism – they do not argue for massive error.

In fact, the ontological, massive-error-entailing claims made by eliminativists tend to arouse suspicion from experimentalists. Stephen Stich, for instance, has long since given up his former eliminativistic ambitions on the grounds that eliminative arguments involve a tacit appeal to descriptive theories of reference (see Stich 1996). It would thus be quite surprising if the arguments pressed by Stich and his colleagues against the use of intuition in philosophical theorizing were of such a form. It is my contention that they are not.

Experimentalist variation arguments, such as the study of Gettier intuitions in Weinberg, Nichols and Stich (2001), do not make claims about the ontological status of entities referred to by philosophically interesting terms. They do not use eliminative scenarios to argue that our positive classificatory judgments about those entities are inevitably erroneous. Instead, they argue that reactions to thought experiments can be influenced by factors which are not relevant to a theory of the phenomenon under study. They then claim that the influence of these factors presents a serious obstacle to the use of thought experiments in the formation of successful philosophical theories. Such arguments by no means rest on the possibility (or actuality) of ‘massive error’ resulting from an eliminative skeptical scenario.

Do such arguments rest on the possibility of ‘massive error’ in some other way? Not obviously. The empirical data reported in Weinberg, Nichols, and Stich (2001) concerns, for example, the tendency of subjects with East Asian backgrounds to ascribe knowledge in Gettier conditions. Given that Western subjects do not tend to ascribe
knowledge in Gettier conditions, something has gone wrong. Perhaps Western and East Asian subjects are referring to different phenomena when they use the word ‘knowledge’ (in which case the relevant question might become “which of these phenomena has more normative force?”). Perhaps they are referring to the same phenomenon, but one group has some false judgments about that phenomenon. Neither of these scenarios requires massive error on either side. The data, at least, are consistent with the absence of massive error – nothing thus far suggests that our knowledge-judgments are on the whole false.

In fact, Weinberg et al. expect that the disagreement found over the Gettier case will not be particularly common.

“We certainly do not mean to suggest that epistemic intuitions are completely malleable or that there are no constraints on the sorts of epistemic intuitions that might be found in different social groups. Indeed, the fact that subjects from all the groups we studied agreed in not classifying beliefs based on ‘special feelings’ as knowledge suggests that there may well be a universal core to ‘folk epistemology’” (Weinberg et al. 2001, 31).

For all the authors have said, most of our judgments about knowledge may be true. This does not straightforwardly falsify the claim that disagreement over particular cases presents a formidable barrier to epistemological theory-building. However, a positive defense of the claim may require some finesse. It remains to be seen exactly how the anti-intuitionist can legitimately generalize from variation findings to a claim that philosophical methodology is broadly suspect, without also providing reasons to doubt the entirety of our classificatory practices. I will attempt to propose a solution in the final section.

4. Philosophy versus everyday cognition
Williamson’s quite reasonable contention that reactions to philosophical thought experiments employ the same cognitive capacities as mundane acts of concept application bars the experimentalist from arguing that the cognitive capacities employed in evaluating thought experiments are generally unreliable. To avoid skepticism, an argument against current philosophical methodology must pursue a different path. What we would like, ideally, is an argument to the effect that these classificatory practices are legitimate and reliable in everyday cognitive activity, while being woefully inadequate in the context of philosophical theory-building.

There is a particular real-life example which demonstrates the possibility of this sort of case quite nicely; Williamson mentions it himself.

“Folk physics… is a theory whose content includes the general principles by which expectations of motion, constancy, and the like are formed online in real time… presumably it is strictly speaking false: although many of its predictions are useful approximations, they are inaccurate in some circumstances; knowledge of the true laws of motion is not already wired into our brains, otherwise physics could be reduced to psychology. Since folk physics is false, it is not known. But the conclusion that no belief formed on the basis of folk physics constitutes knowledge is wildly skeptical” (Williamson 2007, 146).

The particular expectations of motion and other predictions granted to us through use of folk physical principles tend to be true. That’s a good enough reason as any to continue to employ them in our everyday cognition. However, no physicist in his right mind attempts to construct his theories on the basis of the principles of folk physics. For one thing, we don’t have conscious access to the principles upon which our folk physical predictions are made. One could attempt to formulate a set of principles which best fit the pronouncements of folk physics, but this would still be of no use whatsoever to the
study of *physics* – because the principles which best fit folk physical prediction and explanation are *radically false*. \(^{41}\)

The principles by which folk physics operates in fact seem to resemble medieval impetus theory, upon which motion is dependent upon an impetus implanted in a moving object by the cause of the motion; the gradual dissipation of this impetus then explains why, for instance, a thrown object eventually falls to the ground (see McCloskey 1983 for further discussion of the principles of folk physics). Impetus theory, however, is false; the idea that motion requires a force continually sustaining it has been rejected by physicists since the advent of Newtonian mechanics. Thus, though most of our folk-physics-based beliefs are true, folk physics is clearly not a good guide in theory-building – the principles upon which it achieves its reliability are false.

The moral is quite general. Consider a group of ten objects, named a, b, c... j, and two properties, named F and G. Consider a subject whose ‘folk theory’ of those objects and their properties produces the judgments Fa, Fb, Fc ... Fj, Ga, Gb, Gc... Gj. \(^{42}\) In actuality, however, ~Fa and ~Gb – all other judgments are correct. Out of twenty judgments, the subject has made 18 correctly. Suppose, however, that he is a philosopher, and that he is concerned with the nature of F-hood and of G-hood. He might then form some further beliefs about the nature of F-hood and G-hood on the basis of those initial classificatory judgments – he might, for instance, infer that everything (in the toy universe of 10 objects) is F, that everything is G, and that if something is F then it is G. He would be wrong on all counts. The example is simple, but it shows that a certain

---

\(^{41}\) There is of course a further reason not to build one’s physical theories on the pronouncements of folk physics – we have access to a superior source of data, i.e. actual observation of the events about which folk physics generates predictions.

\(^{42}\) The example need not rest on the use of folk theories – insert any belief-forming process you like.
principle – that the general reliability of one’s classificatory judgments directly entails the general success of one’s theory-building – is clearly false.

Let’s return to one of the possibilities suggested earlier to account for the variability in Gettier judgments across cultures. Suppose Western and East Asian subjects are referring to the same phenomenon when they use the word ‘knowledge’ – they are both referring to knowledge. However, since they disagree about whether or not the Gettier case counts as knowledge, one or the other group has a false belief about knowledge. Suppose (probably contrary to fact) that this is the only type of case on which the two groups disagree. Finally, suppose the East Asian subjects are right; the Gettier cases are in fact cases of knowledge.

The percentage of knowledge cases that are also Gettier cases is obviously fairly low. Both groups are highly successful categorizers of knowledge and non-knowledge. But the theories produced by philosophers in each group will differ in philosophically significant ways. The Western philosophers will devote decades of literature to an attempt to formulate new principles which rule out Gettier cases as knowledge; the East Asian philosophers will stick with the classical JTB model. And despite their general reliability as classifiers, the Western philosophers’ theories will simply be false. As in the folk physics case, it’s plausible to conclude that our practices of concept application are just fine for ordinary purposes\textsuperscript{43} – further, unlike the eliminativist, we may conclude

\textsuperscript{43} These ‘ordinary purposes’ also include most uses of concept application in scientific contexts. An experimenter’s report of her observations might, for instance, take the following form – “The samples, when viewed under the microscope, exhibited a blue tint”. In making this report, the experimenter applied the concepts ‘microscope’ and ‘blue’. The possibility of a culture whose reactions to thought experiments involving the words ‘microscope’ and ‘blue’ differed somewhat from our own does not undermine her report. It does, however, present an obstacle to a philosopher interested in analyzing the conditions under which something counts as a microscope, or the conditions under which something counts as blue. As a general rule, you can successfully employ a concept without possessing a correct theory of the referent of that concept. But further, you can successfully employ a concept even if it is impossible to construct such a
that such applications typically result in true beliefs. But it is not unreasonable to look to other methods in the tricky business of theory-building – nor is it unreasonable to question a philosopher who intends to use the Gettier proposition as evidence for his preferred account of knowledge.

The issue here is not that we have reason to believe that every judgment in the knowledge context is likely to be misleading; by hypothesis, the vast majority of such judgments are true. Rather, it is that we have reason to believe that at least some judgments in the context are false – reason that goes beyond a general admission of fallibility. For we have reason to suspect that our concepts are ‘folk physics-y’; that our concepts approximate the categories philosophers are interested in well enough for everyday purposes, but that the principles by which we make these approximations are inadequate. If we are in fact in a position like this, one would be ill-advised to theorize by attempting to form accounts to best fit the entire set of intuitive judgments.

I have argued that general success in sorting particular cases does not entail success in the context of philosophical theory-building. If the studies performed thus far indicate a general trend\footnote{Whether there does exist a general trend of inappropriate variation is of course an empirical question; however, findings thus far make the possibility non-negligible, to say the least.}, at least some of our judgments on any given philosophical phenomenon are likely to be false, and the theory which best captures intuition may therefore be quite different from the theory which best captures reality. To indulge in a bit of speculation, we might further note that some features of the judgments philosophers
focus on suggest that philosophical cases might be just the sort of cases on which a generally reliable yet ‘folk physics-y’ concept-application ability would be likely to err. For one, while philosophically relevant cases like the Gettier scenario are not necessarily sci-fi bizarre (though some are!), they are usually non-prototypical in some regard. For another, the philosophically important cases tend to be the cases which provoke the most disagreement – there is little philosophical discussion devoted to the question of whether or not humans count as conscious beings, for the simple reason that everyone who accepts the notion of consciousness agrees that they do. The list might go on – see Weinberg (2007) for a general discussion of some potential epistemological flaws of philosophical judgments.

Lastly, we might also briefly discuss a second potential account of the differing intuitions in the Weinberg et al. article. It’s entirely possible that Western and East Asian subjects were talking past one another; while Western subjects use the word ‘knowledge’ to refer to one phenomenon, knowledge1, East Asian subjects use the word to refer to a different phenomenon, knowledge2. Further, no eliminative scenario obtains; there really are things that fall under the category knowledge1, and there are also things that fall under the category knowledge2. Both Western and East Asian subjects might then have generally true beliefs about their respective phenomena – indeed, they might even be perfectly reliable in their ability to classify cases as instances of those phenomena – but important theoretical questions might still remain. We might wonder which of these phenomena has more normative force; or, we might wonder which plays some important theoretical role, such as being the most general factive mental state. Those questions are not answered by a mere ability to correctly sort cases.
It’s time to summarize. It is plausible that Evidence Neutrality is not a genuine constraint; there are some views which we are not required to countenance when we are in the business of tallying our evidence. In particular, views like radical skepticism about perception leave us with an evidential base which is insufficient for our philosophical purposes. Any view that narrows our evidential base to such an extreme degree may be justly ignored. Williamson claims that critics of intuition force us to radically ‘psychologize’ our evidence, thereby unduly narrowing our evidential base – the critics themselves, however, do not take themselves to be committed to such restriction.

One reason to think that critics of intuition might be committed to pervasive skepticism is the plausible claim that ‘intuitions’ are really just instances of our ordinary cognitive capacities for counterfactual evaluation and for concept application. If one questions the use of intuition, one questions the ubiquitous practice of applying concepts in judgment – if one argues that intuitions are prone to massive error, one argues that we are in massive error across the board. Critiques along the lines of Weinberg et al., however, do not rest on claims of massive error. They are, in fact, consistent with the truth of the vast majority of our classificatory judgments. Part of their force comes from the fact that it does not take much error to seriously disrupt successful theory-building; though one’s classificatory judgments may be largely correct, the theoretical principles that best fit such judgments may be radically false. The claim that our capacity for concept application produces reliable general-purpose judgments, while being distinctly unsuited to the goals of philosophical theory-building, is not only consistent, but fairly plausible. The general falsity of folk physics is uncontroversial, yet judgments resulting from its use are typically true; it is not unlikely that other mental capacities are similarly
‘good enough for government work’ even though they operate on false theoretical principles.

Any argument of this sort against the philosophical use of intuition involves no threat of general skepticism, and cannot be legitimately set aside on grounds of the falsity of Evidence Neutrality. This is not to claim that no thought experiment is ever relevant to the assessment of philosophical theories; as Williamson notes, thought experiments have sometimes been decisive in scientific debates, and there is no reason to doubt that the same will be true of philosophy. But the pervasive use of reactions to thought experiments as the primary data for philosophical theorizing can, and should, be put under serious scrutiny.
Chapter 4

Does Intuition Suffice?

In a series of papers (Bealer 1987, 1992, 1996, 1998, 2000), George Bealer has defended a conception of the role of intuition in philosophy which epitomizes what might be called the ‘traditionalist’ view. According to Bealer, intuition is a genuine source of evidence, the use of which is indispensible to the methodology of philosophy. Intuition is not only indispensible to the task of philosophy, but adequate, as well; philosophical questions can in principle be answered without any substantive help from the sciences. Thus, philosophy is an autonomous discipline. In addition to its autonomy, Bealer claims, philosophy possesses authority – the answers philosophy gives to philosophical questions generally cannot be overturned by conflicting answers provided by science. Though these two theses present an appealing outlook on the discipline of philosophy, it is my contention that they are false.

Bealer offers two primary arguments in support of the theses of autonomy and authority. The first, the ‘Argument from Evidence’, begins by providing reasons for accepting the role of intuitions as a reliable source of evidence. The argument then endeavors to show that the best explanation of intuition’s evidential status is that intuitions possess a strong modal tie to the truth. The second argument, the ‘Argument from Concepts’, provides an account of concept possession which, if true, would provide intuition with the postulated strong modal tie – thereby grounding claims of autonomy and authority. There are, however, serious problems with both arguments, as this paper will aim to show.
Against the Argument from Evidence, I will first suggest that Bealer’s arguments for the evidencehood of intuition prove much less than they aim to – at best, they show that certain very basic appeals to intuition cannot be eliminated from inquiry. I will then claim that, even should we grant that intuition is as basic a source of evidence as (e.g.) phenomenal experience, this would not imply that philosophy possesses either autonomy or authority. Against the Argument from Concepts, I will propose that Bealer’s account of concept possession is at best unmotivated, and at worst deeply implausible. Though Bealer argues that we can achieve complete understanding of philosophical concepts entirely a priori, I will argue that full understanding of such concepts may plausibly depend on our knowledge of contingent, empirical facts. I will conclude with a general discussion of the links between philosophy and the sciences.

1. The Argument from Evidence

As noted above, Bealer’s overall defense of the autonomy and authority of philosophy has two parts; the Argument from Evidence, and the Argument from Concepts. Though Bealer sometimes presents the two arguments as largely independent, they are in fact intimately related; the Argument from Concepts attempts to provide the actual mechanism underlying the ‘strong modal tie’ to truth hypothesized by the Argument from Evidence. As such, I will present them as forming a unified account of the role of intuition and of the autonomy and authority of philosophy. We’ll begin, in this section, with an examination of the Argument from Evidence.
The Argument from Evidence consists of several steps. The first is to motivate the claim that intuitions are, in fact, evidence. Bealer claims that reliance on intuition forms part of our ‘standard justificatory procedure’ – that is to say, we standardly employ intuitions as evidence for or against various beliefs, theories, philosophical hypotheses, and so on. Of course, the mere fact that intuitions are employed as evidence does not suffice to demonstrate that they ought to be treated in this way. However, Bealer argues that any attempt to formulate an alternative justificatory procedure which eschews reliance on intuition would lead to epistemic self-defeat. In his (1992), Bealer in fact provides three distinct arguments in support of this accusation of self-defeat; it will be valuable to examine each.

Bealer’s first argument against the possibility of modifying our standard justificatory procedure to exclude intuitions as a source of evidence involves the claim that any such modification would have to be motivated from within the standard justificatory procedure itself. This is not, I take it, a claim that the use of an alternative justificatory procedure cannot count as justified simpliciter unless it is deemed justified by the lights of the standard justificatory procedure – after all, the standard justificatory procedure is just that procedure which we in fact standardly employ, and it is surely at least possible that the procedure could be wholly misguided. Were our standard justificatory procedure to be wholly misguided, it might deem unjustified even an alternate procedure which was, in fact, epistemically optimal.

Instead, I take it that Bealer is making a more practical claim – adherents of the standard justificatory procedure cannot (as a matter of psychological or dialectical fact) be persuaded to adopt an alternative procedure unless provided reasons which conform

45 One assumes Bealer intends this ‘we’ to cover Western philosophers, though this is not specified.
with the procedure they currently use. This is merely to say that there must exist some ‘common ground’ between the competing procedures; for instance, if both the standard and the alternate procedures place high epistemic value on consistency, and it can be shown that the standard procedure (but not the alternative) leads to inconsistency, this might be used to convince adherents of the standard procedure of the superiority of the alternative.

Ideally, then, proponents of an alternate procedure will want to attempt a critique which employs principles found within the standard procedure. Bealer claims that our standard justificatory procedure does, in fact, provide us with the means to critique currently accepted sources of evidence from within, and potentially reject them if they are found wanting. We do so by reference to several factors, including intra-personal consistency, inter-personal corroboration, and external confirmation. None of these three considerations, Bealer argues, undermines the evidential status of intuition.

According to Bealer, the intuitions of particular individuals are generally quite consistent across time. Further, our intuitions are generally shared by others; though there are bound to be cases of disagreement, Bealer claims that there is ‘impressive corroboration’ among our more basic intuitions. Finally, our intuitions are rarely, if ever, disconfirmed by experience. Our standard procedure for critiquing evidential sources therefore leaves intuition unscathed – or so the argument goes.

Unfortunately for Bealer, recent empirical studies – in fairness, most conducted several years after the formulation of Bealer’s argument – have thrown these claims into serious doubt. There is, for example, substantial evidence that an individual’s moral judgments may be quite inconsistent across time. The well-known ‘Asian disease’ study
by Tversky and Kahneman (1981) demonstrated the effects of framing on moral
judgment, while Haidt, Koller and Dias (1993) and Wheatley and Haidt (2005) have
demonstrated strong effects of disgust. Similar findings exist for epistemological
judgments; Swain, Alexander and Weinberg (2008) have found evidence of order effects
in response to Truetemp cases. Studies on cross-cultural variation in intuition, as in
Weinberg, Nichols and Stich (2001) and Machery, Mallon, Nichols and Stich (2004),
indicate that intuition lacks much by way of interpersonal corroboration, as well.

The question of intuition’s capacity for external confirmation is somewhat more
complicated. Bealer writes that “intuition is seldom, if ever, disconfirmed by our
experiences and observations. The primary reason is that the contents of our intuitions…
are by and large independent of the contents of our observations and experiences” (Bealer
1992, 110). A main thesis of the second half of this chapter will be that our intuitions
can, and not infrequently have been, disconfirmed by empirical discovery. For now,
however, it is worth noting that at least one critic of intuition would take Bealer’s claim
of immunity to disconfirmation to support the exact opposite of his conclusion about the
evidential pedigree of intuition. Cummins (1998) argues that, if we are to be justified in
employing their deliverances, our evidential sources must be capable of being calibrated
– i.e., we must be able to check their outputs against those of an independent source. If
intuition cannot be calibrated in this way – if there is no non-intuitive access to the targets
of intuition – then this is plausibly an epistemological flaw, rather than an asset.

There are, of course, responses Bealer might make to each of the arguments just
discussed; he might, for instance, side with Sosa (2007) in claiming that the variation
demonstrated by the studies mentioned above is no worse than the variation to which
vision is occasionally susceptible. Or he might suggest that the responses elicited by such studies are not intuitions of the relevant sort.\textsuperscript{46} Debate about the significance of the findings of experimental philosophers is still lively; and I have little new to add to those debates here. I will simply make the following observation: Bealer’s argument for the immunity of intuition to standard evidential criticism is open to serious doubt.

Bealer’s second argument against the possibility of rejecting intuition is somewhat narrower in scope. Its target is the Quinean empiricist, whom Bealer takes to endorse the following epistemological principle: “a theory is justified for a person if and only if it is, or belongs to, the simplest comprehensive theory that explains all, or most, of the person’s experiences and/or observations” (Bealer 1992, 118). Bealer also takes the Quinean empiricist to be committed to the claim that this ‘simplest comprehensive theory’ is in fact identical to the simplest regimented formulation of the natural sciences, plus the logic and mathematics required by them. The trouble, according to Bealer, is that the simplest regimented formulation of the natural sciences does not contain the Quinean’s original epistemological principle. The Quinean’s principle is an epistemological one – it contains the terms ‘justified’, ‘theory’, etc. However, Bealer claims that such terms do not occur in the primitive vocabulary of the simplest regimented formulation of the natural sciences. Because the Quinean cannot advert to

\textsuperscript{46} I expect that the ‘wrong type of intuition’ reply would in fact be Bealer’s most likely response – in a piece written before all the studies mentioned save Tversky and Kahneman’s, Bealer claims that “many philosophers believe that the empirical findings of cognitive psychologists such as Wason, Johnson-Laird, Rosch, Nisbett, Kahneman, and Tversky cast doubt on [intuitions’] epistemic worth. But, in fact, although these studies bear on ‘intuition’ in an indiscriminate use of the term, they evidently tell us little about the notion of intuition we have been discussing” (Bealer 1998, 213). At least in part, Bealer takes the cognitive psychologists’ notion to be ‘indiscriminate’ in that they have not explicitly tested for the phenomenological features Bealer takes to be definitive of intuition. I will not attempt to criticize this claim here, but see my arguments in Chapter 2 for some reasons to be wary of Bealer’s phenomenological definition of intuition.
such terms, the Quinean’s epistemological theory is not justified according to the
Quinean’s epistemological theory. The theory is thus self-defeating.

I will here make no claims as to the validity of this argument, for I take it that the
argument simply does not target the majority of current participants in the intuition
debate. As far as I can tell, there is nothing inherent in the anti-intuitionist position that
requires one to endorse the Quinean’s epistemological principle; certainly, such a
principle has played little part in argumentation within the intuition literature of the past
fifteen years. Further, should an anti-intuitionist feel drawn to something like the
Quinean epistemological principle Bealer describes, I do not see any reason why she
could not grant epistemology a position next to logic and mathematics in the formulation
of the ‘simplest comprehensive theory’. Such a move would not be particularly ad hoc –
plausibly, certain basic epistemological principles are required to successfully conduct
scientific inquiry, just as logic and mathematics are. Granted, there remains the question
of whether these epistemological principles can be defended without use of intuition –
but this is just the question raised by Bealer’s final self-defeat argument. The argument
against the Quinean empiricist, then, to my mind adds no new difficulties.

Bealer’s final argument for self-defeat relies on the following observation: in
order to formulate any successful justificatory procedure, one must invoke certain basic
epistemic classifications, or ‘starting points’. These starting point classifications involve
judgments regarding what does and what does not count as an experience, an observation,
an explanation, a valid argument, and so on. If the anti-intuitionist is to formulate an
alternative to our current standard justificatory procedure, she will have to invoke at least

47 To be clear, Bealer doesn’t take the Quinean principle to be required by the anti-intuitionist, either;
indeed, in his (2000) he mentions that non-Quineans will likely find the other self-defeat arguments more
persuasive.
some basic epistemic classifications, and thus rely on at least some ‘starting point’
judgments. In order for her claim that intuitions do not count as evidence to make sense,
for example, the anti-intuitionist must be able to sort evidence from non-evidence.

The trouble, Bealer claims, is that starting point judgments are invariably arrived
at via intuition. Thus, the anti-intuitionist herself will have to depend on intuition in
order to formulate her alternate justificatory procedure. In practice, then, the anti-
intuitionist cannot avoid the use of intuitions as evidence; it is impossible to defend the
rejection of intuition without appeal to intuition. Thus, any argument against the use of
intuition is bound to be self-defeating – for any such argument will advocate rejection of
an epistemological resource which is essential to its formulation.

Bealer predicts that the anti-intuitionist will respond to this worry by
distinguishing between the \textit{source} of a belief and the \textit{current grounds} for holding that
belief – that is to say, she will invoke a distinction between the ‘context of discovery’ and
the ‘context of justification’. Although the anti-intuitionist will likely have to admit that
she initially formulated her beliefs about basic epistemic classifications via intuition, she
may claim that she does not rely on those intuitions for her current justification. Bealer
responds to this anticipated move by claiming that, although the strategy is legitimate in
principle, it leaves the anti-intuitionist caught in a fatal dilemma.

To begin, clearly it is either the case that the anti-intuitionist’s intuitions about
basic epistemic classifications are reliable, or that they are not. Assume they are \textit{not}
reliable; thus, the intuition-based judgments the anti-intuitionist forms about what counts
as a theory, as justification, etc., are prone to error. Bealer argues that this error will be
reflected in any theories based on those judgments. As a result, it is irrelevant whether or
not the anti-intuitionist can discover alternate, non-intuition based justifications for the judgments she initially based on intuition. The judgments will, by hypothesis, be error-ridden. Any theory that results from their use will inherit this unreliability; thus, the theories formulated by the anti-intuitionist will be unreliable.

Bealer does note that, during the process of theory formation, one can often identify and eliminate erroneous judgments. During the development of the various sciences, for example, we have been able to identify contexts in which our perceptual faculties (upon which science depends) err. But he notes that in the case of perception, it is precisely because our perceptions are on the whole reliable that we are able to make the required corrections.

“Suppose that a person suffers from an astigmatism, making his visual observations of shape and length prone to error. Again, it is plausible that, upon formulating a comprehensive and systematic theory on the basis of all his observations, including the largely reliable observations provided by his other senses and by his largely reliable visual observations of color, continuity, contiguity, and other topological properties, the person will be able to spot and eliminate these errors about shape and length” (Bealer 1992, 106).

By contrast, if perception were on the whole unreliable, such corrections would be impossible. If intuition is unreliable, Bealer argues, then error at the theoretical level will not be remediable.

To take the other horn of the dilemma, assume that the anti-intuitionist’s intuitions about basic epistemic classifications are reliable. If this is the case, Bealer argues, then whatever it is that makes these basic epistemic intuitions reliable will also make our intuitions about what counts as prima facie evidence reliable. Further, “we have a wealth of concrete-case intuitions to the effect that intuitions are prima facie evidence” (Bealer 1992, 107); since these intuitions are by hypothesis reliable, “it would follow that intuitions are in fact prima facie evidence” (Bealer 1992, 107). Thus, the anti-intuitionist would be advocating rejection of a legitimate source of evidence.
“If intuitions are prima facie evidence, then the sort of overall theory that empiricists would formulate (after excluding intuitions as prima facie evidence) would be highly unreliable (notably, on such matters as modality, definition, property identity, evidence, and justification” (Bealer 1992, 107).

The anti-intuitionist, then, ends up with an unreliable theory either way.

There are, I believe, multiple places at which one can question this argument.

With regard to the second horn of the dilemma, why should we think that whatever makes us reliable identifiers of observations or of valid arguments should make us reliable identifiers of prima facie evidence? Indeed, why should we think that what makes us reliable identifiers of observations should be the same as what makes us reliable identifiers of valid arguments? As I argue in chapter 2, it is quite plausible that our intuitive capacities are heterogeneous – if this is right, then the inference from ‘humans are reliable identifiers of observations’ to ‘humans are reliable identifiers of prima facie evidence’ requires argument.

Worse, though, it’s not even clear that the argument succeeds even if we take intuition to be homogenous. Bealer moves from the assumption that humans are reliable at basic epistemic classification to the assumption that humans are reliable at identifying prima facie evidence. This is essentially analogous to the assumption that, since humans are reliable at visual identification of objects, we are reliable at visual identification of cleverly disguised mules. The reliability of a given faculty does not at all entail the reliability of any particular subset of judgments produced by that faculty.

---

48 There is room for a bit of confusion over use of the phrase ‘prima facie evidence’. On one reading, prima facie evidence consists of whatever appears, prima facie, to be evidence. On this reading, it is indeed implausible that we should not be reliable identifiers of prima facie evidence – after all, we surely have access to the way in which things appear to us. However, this is not the reading that I take Bealer to intend. For something to be prima facie evidence is (roughly) for it to belong to a type that is standardly evidential, though this evidential status might be defeated by competing considerations. Thus, vision is prima facie evidence, even though not all instances of vision are evidence – vision under poor lighting conditions, for example, might fail to be evidence. Clairvoyance is not prima facie evidence, even if it happens to appear to me to be evidence at first glance.
Though the on-balance reliability of intuitive epistemic classification might lend plausibility to the claim that humans are \textit{probably} reliable at evidence-identification, plausibility is not enough for the sort of argument Bealer is here attempting to press. Bealer aims to claim that arguments against the use of intuitions as evidence are self-defeating, and that they therefore \textit{cannot} succeed. At best, though, an argument that it is plausible that we are reliable identifiers of prima facie evidence puts the burden of proof on the anti-intuitionist to show that those intuitions are misleading. This is a far cry from self-defeat.

Finally, even putting this problem aside, granting both human reliability with regard to identification of prima facie evidence and the claim that we intuit that intuitions are prima facie evidence, it simply does not follow that intuitions are \textit{in fact} prima facie evidence. No such conclusion can follow unless intuition is \textit{perfectly} reliable. And, quite reasonably, Bealer admits that it is not; intuition is fallible. It is thus perfectly consistent to claim that, while we are on the whole reliable identifiers of prima facie evidence, our intuitions have steered us wrong in this particular instance. Again, though the reliability of evidence-identification would lend plausibility to the claim that our intuition that intuitions are prima facie evidence is accurate, this is radically insufficient for the purposes of Bealer’s argument that anti-intuitionists face self-defeat.

Let’s now examine the first horn of Bealer’s dilemma, and consider what follows if intuitions about basic epistemic classifications are assumed to be \textit{not} reliable. Bealer claims that error correction would be impossible were intuition not on the whole reliable. But he seems to assume that the evidential source that enables error correction must be the very same as the evidential source which generated the faulty judgments. Why
shouldn’t we employ our other cognitive resources in order to correct errors generated by our initial intuitions about what counts as justified, or as explanatory? Even if our basic epistemic classifications are unreliable, we presumably possess an enormous body of beliefs, generated by all sorts of cognitive mechanisms, which are reliably formed. Why shouldn’t this body be sufficient for the construction of theoretical principles which would enable us to amend errors in some of our more unreliable classification intuitions?

The possibility of heterogeneity is relevant here, as well. Imagine that most types of basic epistemic classification intuitions are unreliable. Given a heterogeneous view of intuition, this does not entail the unreliability of logical intuitions, mathematical intuitions, or even other types of classificatory intuition. Though Bealer claims that error in basic epistemic intuitions will not be remediable unless ‘intuition’ (in the broad sense) is reliable, this claim is questionable when intuition is viewed heterogeneously. If our mathematical and logical intuitions are sufficiently reliable, and if we have a sufficient base of reliably-formed beliefs in other domains, we may well be able to correct errors generated by reliance on unreliable epistemic classification intuitions.

Finally, with regard to the argument more generally, the overall strategy of attempting to prove intuition’s evidential status by demonstrating that radical empiricist positions face self-defeat is puzzling. It is unclear why we should accept a move from the impossibility of rejecting basic epistemic intuitions (due to looming self-defeat) to a conclusion regarding the evidential status of intuition as a whole. Again, the possibility of heterogeneity should bring this move under serious scrutiny. Assume that at least some types of basic epistemic classification intuitions are reliable. On the ‘heterogeneous mechanisms’ view of intuition I urged in chapter 2, this entails nothing whatsoever about,
e.g., the reliability of moral intuition or of intuitions about consciousness. The anti-intuitionist can, then, successfully formulate a justificatory procedure which rejects the use of at least some types of ‘intuition’ as evidence.

Further, though – and more importantly, for the purposes of this paper – the claim that intuition is a reliable source of evidence does not yet suffice to demonstrate the theses of autonomy and authority. Consider an analogy. Suppose there are two methods of estimating the age of certain archaeological samples. The first is very reliable, but can only provide estimates for objects less than 100,000 years of age. The second method is much less reliable than the first (though perhaps still reasonably reliable), but can provide estimates for objects of up to 110,000 years of age. Clearly, there are bound to be certain objects for which use of the second method is the only option; however, this in no way implies that the second method is an ‘autonomous’ or ‘authoritative’ method. By hypothesis, the first method will be authoritative over the second in the majority of cases. And the second method is only ‘autonomous’ in that it can provide more-or-less reliable estimates should the first method be unavailable.

While I will refrain from claiming that this analogy closely represents the actual differences between philosophical and empirical methodologies, the moral should be sufficiently clear. An argument that intuition is indispensable in certain situations – indeed, even an argument that intuition is on the whole reliable – simply does not entail that ‘the central questions of philosophy…can in principle be answered by philosophical investigation and argument without relying substantively on the sciences’ (Bealer 1998, 201). Nor does it entail that the answers provided by intuition are authoritative over those provided by the sciences. Something more will be needed.
In order to generate a complete argument for the autonomy and authority of philosophy, the arguments just reviewed must be coupled with the second half of the Argument from Evidence. Recall that Bealer has concluded that intuitions must be evidence. Bealer next proposes that the best explanation of intuition’s evidential status will advert to intuition’s reliability - the tendency of its deliverances to be true. Reliability is not, according to Bealer, a general requirement on something’s counting as evidence; it is, however, a requirement for basic epistemic sources like intuition. Indeed, something is a basic source of evidence if and only if it possesses an appropriate sort of reliable tie to the truth.

Why think that intuition is a basic source of evidence? The argument is strikingly simple: intuitively, intuition is a basic source of evidence. Though this may sound circular, it is not – so long as it has already been concluded that intuitions are evidence, intuitions may be used to back the claim that intuition is a basic source of evidence.\(^49\) I will not challenge this stage of the argument, though I must admit to lacking the intuition in question myself.\(^50\)

What sort of reliable tie to the truth characterizes basic sources of evidence? Whatever its nature, it cannot be merely contingent: if contingent reliability were sufficient for basic evidencehood, a being who just so happened to possess contingently reliable clairvoyance might count as thereby possessing basic evidence. Even nomologically necessary reliability is not strong enough, for similar reasons. Instead,

\(^{49}\) It is true that, again, the overall reliability of intuition does not entail that our intuition that intuition is a basic source of evidence is correct. But presumably it would lend a fair amount of plausibility to such a claim. Unlike the case of the earlier, self-defeat arguments, plausibility is arguably sufficient for Bealer’s current purposes.

\(^{50}\) Bealer as a brief secondary argument for the basicness of intuition’s evidencehood. However, this argument assumes that his opponent endorses the Quinean epistemological principle discussed earlier, and is subject to the same concerns I raised during that discussion.
Bealer claims, basic sources of evidence must have a *strong modal tie* to the truth; that is to say, “necessarily, they are always true, mostly true, probably true, or normally true” (Bealer 1987, 31). Given what we know of the fallibility of intuition, however, the modal tie that holds in their case must not be so strong as to imply that error is impossible; the explanation of intuition’s evidencehood must invoke a modal tie that holds relative to suitably good cognitive conditions, and that holds only most of the time.

So concludes the Argument from Evidence. Intuitions must be evidence, because to argue otherwise is to face self-defeat. Intuition tells us that intuitions are *basic* evidence. Basic evidence must possess a strong modal tie to the truth; ergo, intuition must possess a strong modal tie to the truth. Now that the argument is concluded, the link to autonomy and authority begins to seem somewhat clearer. A source of evidence whose deliverances are *necessarily* mostly true has, at least on one dimension, epistemological credentials which are stronger than those of scientific inquiry. For, it is *not* the case that empirical methods of investigation provide us with hypotheses that are necessarily mostly true.

However, it is not obvious to me that the link to autonomy and authority is, on closer examination, substantially clearer after all. If Bealer’s argument succeeds, our intuitive judgments are *necessarily* mostly true (in good cognitive conditions). But the importance of this particular modal characteristic is dubious. The deliverances of

---

51 I admit to substantial puzzlement at portions of Bealer’s argument here. Bealer claims that, for basic evidential sources, possession of a contingent or even nomologically necessary tie to truth is insufficient – a ‘strong’ modal tie must be present, instead. Yet other obvious candidates for ‘basic’ evidential sources seem to possess only contingent ties to truth. Surely it is contingent that perception is reliable – at least insofar as evil demon scenarios are genuine possibilities. Bealer in fact seems to discount ‘observational beliefs’ as basic evidence, opting instead to recognize only ‘phenomenal experience’ – I assume this consists of beliefs of the form “it appears as though a red patch is before me”, or some such. Leaving aside the question of whether such beliefs in fact possess a strong modal tie to the truth, I will simply note that this strikes me as an abnormally narrow conception of basic evidence. I begin to suspect that my ability to intuitively grasp the boundaries of basic evidence is not as reliable as Bealer might have hoped.
scientific inquiry are not of necessity mostly true, but they could quite plausibly be, 
contingently, more frequently true than intuition. Or, they could turn out to be generally 
true in circumstances where intuition (though necessarily on the whole reliable) 
frequently errs. Should that be the case, pursuit of an ‘autonomous’ philosophical 
methodology would be wrong-headed, as would claims to universal authority for intuitive 
answers to philosophical questions. Ironically, at this point the status of the autonomy 
and authority of traditional philosophical methodology may be an empirical question – 
just as the anti-intuitionist would urge.

2. The Argument from Concepts

Let’s now move to an examination of Bealer’s Argument from Concepts. The 
Argument from Concepts attempts to provide an account of concept possession which 
would provide the ‘strong modal tie’ which Bealer has concluded that intuition must 
possess. There are, Bealer claims, two standard ways in which we might speak of a 
person as possessing a concept – she might possess it either nominally, or else 
determinately. A subject possesses a concept nominally iff she can have propositional 
attitudes towards propositions that have that concept as part of their content. A subject 
possesses a concept determinately iff she possesses it nominally, and in addition “does 
not do this with misunderstanding or incomplete understanding” (Bealer 2000, 11). It is, 
Bealer claims, at least possible for the central concepts of philosophy to be possessed 
determinately by some cognitive agent. Thus, it is at least possible that our intuitions are 
necessarily reliable sources of evidence for philosophical hypotheses. If they are, Bealer
claims, then we need not turn to science to answer philosophical questions - intuition provides philosophy with autonomy. And since “no such necessity ever holds for science” (Bealer 1996, 131), the possibility of determinate concept possession means that it is at least possible that the authority of philosophical intuition is greater than the authority of the sciences. As noted in the previous section, it is unclear why modal strength should be relevant to the autonomy or authority of philosophy. However, a full characterization of the nature of our possession of philosophical concepts might yet reveal the source of these purported characteristics.

It is worth noting that the claim Bealer has made appears to be a weak one – it is merely possible that we might achieve autonomous and authoritative philosophical knowledge. However, Bealer suggests that we in fact approximate this cognitive ideal, and that we can, through sufficient a priori reflection, eventually obtain answers to most philosophical questions. Thus, in the actual world, philosophers can in principle proceed without input from the sciences.

In the remainder of this section, I will attempt to show that, although Bealer’s notion of determinate concept possession is coherent, it is not at all obvious that any actual cognitive agent in fact possesses any of her philosophical concepts determinately. Worse, without recourse to empirical investigation of the relevant philosophical kinds, it is unclear how we might pursue determinate concept possession, on the assumption that we currently lack it. If this is in fact the cognitive situation in which we find ourselves, the mere metaphysical possibility of autonomous and authoritative knowledge holds little comfort.
An informal gloss on the notion of determinate concept possession has already been given – to possess a concept determinately is to possess it nominally, and in addition to do so without misunderstanding. The way in which this determinate concept possession purportedly generates necessary reliability can be illustrated by the following example, which Bealer gives in several papers (1987, 1998, 2000). Imagine a woman introduces, through use, a term ‘multigon’; she applies it to various closed figures but has not, as of yet, applied it to or withheld it from triangles or rectangles. Suppose her term expresses a definite concept and that she determinately possesses this concept. Bealer argues that it seems intuitive that when she considers triangles and rectangles (in suitably good cognitive conditions – i.e., she is sufficiently attentive, intelligent, and so forth) she will apply ‘multigon’ to them iff the property of being a multigon is the property of being a polygon, and she will withhold ‘multigon’ from them iff the property of being a multigon is the property of being a polygon with five or more sides. This could fail to be the case only if either she did not in fact possess the concept determinately, or she was not in suitably good cognitive conditions. Thus, her intuitions have a strong modal tie to the truth.

The example is in some ways an odd one. The term ‘multigon’ is introduced by a single individual, and (so far as one can tell from the example) employed by a single individual. Rather than being a term belonging to a community of language users, the term ‘multigon’ is an element in a single speaker’s idiolect. Though it does perhaps seem intuitive that the woman will apply the term ‘multigon’ to triangles and rectangles if and only if triangles and rectangles are in fact multigons, it is much less clear whether she would possess this disposition had she acquired the term from other members of a
language community. Perhaps Bealer only intends to illustrate what he would take to be a central case of determinate concept possession, rather than convince us that such determinate possession is commonplace. Unfortunately, when we put aside the intuitive appeal of the multigon case, the informal notion of determinate concept possession as ‘possessing a concept without misunderstanding’ no longer seems particularly enlightening.

As it stands, ‘determinate concept possession’ is not really an explanation of (or an argument for) Bealer’s proposed strong modal tie, but a rephrasal of it. One possesses a concept F determinately iff one does so without misunderstanding. This is essentially to say that one possesses F determinately iff one is disposed to judge that a thing counts as an F just in case that thing does count as an F. Indeed, Bealer even provides the following characterization – “x determinately possesses the concept of being a multigon iff: x would have the intuition that it is possible for a triangle or rectangle to be a multigon iff it is true that it is possible for a triangle or rectangle to be a multigon” (Bealer 2000, 15; emphasis original). Thus far, then, we have not been given much reason to think that our intuitions in fact possess a strong modal tie to the truth. All we have is an indication of what our concepts would have to be like in order for the strong modal tie to hold.

Indeed, there is an obvious set of cases for which we would expect this sort of modal tie not to hold – natural kind terms. Were the average English speaker to possess, say, the concept elm determinately in the sense just provided, then it seems she should be disposed to intuit, given a description of a possible tree, that such a tree counts as an elm iff the described tree in fact counts as an elm. Experienced dendrologists aside, this is
rarely the case. In other words, the informal characterization of determinate concept possession just given seems to conflict with the Kripke/Putnam view that natural kind property-identity statements are necessary yet a posteriori – a view which Bealer terms ‘scientific essentialism’. Scientific essentialism is relatively uncontroversial for terms like ‘water’ – and therefore, it seems uncontroversial that no strong modal tie holds in the case of our water intuitions. It’s just not the case that, through a priori reflection alone, I could come to possess the concept ‘water’ so determinately that I would intuit that x counts as water just in case x is H₂O. My water intuitions simply do not possess necessary reliability.

Bealer’s response to this worry is to claim that scientific essentialism simply does not apply to the vast majority of philosophically relevant concepts. Unlike terms like ‘water’, philosophical terms are ‘semantically stable’. If a term is semantically stable, the contingent features of the speaker’s environment do not contribute to its meaning. Thus, an allegedly semantically stable term like ‘knowledge’ is not ‘twin-earthable’; all of my doppelgängers mean the same thing as I do when they make ascriptions of knowledge. There are, Bealer admits, some philosophically relevant terms like ‘time’, ‘space’, ‘cause’, and ‘probable’ which may be semantically unstable. However, Bealer (1996) proposes that these terms have ‘generic’ uses which are semantically stable, as evidenced by phrases like “Euclidian space is a possible kind of space”; it is these generic uses which underlie the general philosophy of e.g. space and time.

This move is, I think, open to serious question. Though the terms ‘space’ and ‘time’ may have the generic uses Bealer describes, it is not at all clear that most metaphysicians and philosophers of science take themselves to be investigating these
generic notions. It is also less than clear that terms like ‘cause’ and ‘probable’ have a
generic use. Worse, the list of semantically unstable philosophical terms is potentially
much larger than Bealer supposes. As Lycan (1996) has noted, it is plausible that many
terms investigated by the philosophy of mind, such as ‘consciousness’, are not
semantically stable (and, Lycan claims, do not have generic uses). It is fairly likely that
the same holds for ‘color’. The thesis that ‘knowledge’ denotes a natural kind, and thus
presumably possesses semantic instability, has been defended by at least one prominent
epistemologist (Kornblith 2002). Similar positions with regard to even moral terms like
‘good’ and ‘ought’ have been held by, e.g., Richard Boyd (1988).

It is, admittedly, unclear whose court the ball is currently on. Some philosophers
have maintained views on the central concepts of philosophy which would support an
attribution of semantic stability to the relevant terms; some philosophers have defended
views which are less compatible with semantic stability. Bealer does provide at least one
argument in support of the claim that the majority of philosophical terms possess
semantic stability; in order to evaluate it, however, we must first examine Bealer’s
response to the worry of scientific essentialism in more detail.

Though Bealer admits that we do not possess the ability to determine the
extensions of natural kind terms a priori, he claims that we do rely on intuitions
regarding e.g. twin-earth scenarios as evidence for the truth of scientific essentialism
itself; thus, we do have at least some reliable intuitions about the proper application of
natural kind terms. This indicates that there is at least some sense in which we can be
said to possess natural kind concepts determinately. Specifically, Bealer argues that we
possess intuitions which enable us to understand that, for example, water is a
‘compositional’ kind – that what is common to all samples of water is their possession of a certain compositional structure. Bealer uses the term ‘categorial content’ to express the dimension of our concepts which determines the fundamental categories (compositional kind, functional kind, etc.) to which those concepts belong. Insofar as we have determinate understanding of the categorial content of our natural kind concepts, we can be said to understand those concepts determinately.

This determinate understanding gives us some degree of a priori insight into the natures of natural kinds. Due to our determinate understanding of the categorial content of our water concept, we can intuit that if water is composed of H₂O on earth, then a twin earth containing only XYZ would contain no water. Other semantically unstable concepts work in similar fashion; our intuitions enable us to determine that tiger-like beings found on an alien world would not be tigers, for (presumably) they would be members of a different ‘evolutionary’ kind. It is, of course, these sorts of intuitions that we rely on in our responses to twin-earth style thought experiments – and it is these intuitions which provide us with the primary evidence that terms like ‘water’ and ‘tiger’ are semantically unstable.

As for philosophical concepts, Bealer’s view on the a priori accessibility of the categorial content of natural kind concepts provides him with the resources to defend his claim that most terms central to philosophical questions possess semantic stability. Bealer argues that it is intuition that allows us to determine when a term is and is not semantically stable – it is, after all, philosophers’ intuitions about twin-earth style scenarios which led us to conclude that terms like ‘water’ are semantically unstable. More precisely, it is our a priori understanding that water is a compositional kind that
allows us to infer that the actual physical composition of the watery stuff of our acquaintance is determinative of the identity conditions for water. Had ‘water’ picked out a ‘macroscopic’ kind – that is to say, a kind individuated by its macroscopic features, such as being wet, clear and drinkable – we would have intuited that ‘water’ refers to watery stuff in every possible world, regardless of its composition. We would have, in other words, intuited that ‘water’ is not twin-earthable.

The fact of the matter, Bealer claims, is that we lack ‘twin-earth’ intuitions for most philosophical concepts. We do not, for example, intuit that there could be a group in an epistemic situation identical to our own whose term ‘pain’ picked out something other than pain. Due to our a priori access to the categorical content of our concepts, it follows that most philosophical terms are, in fact, not twin-earthable. Therefore, philosophical concepts must be by-and-large stable – the contingent features of the world we inhabit do not contribute to their meanings.

This line of argument does, of course, rely on the assumption that our intuitions on twin-earth cases are reliable. In one sense, this is unobjectionable, given that Bealer takes himself to have already proven the reliability of intuition in his Argument from Evidence. On the other hand, anyone unconvinced by the Argument from Evidence need not take twin-earth intuitions to provide the final word on semantic stability. Depending on one’s views on reference, though, there may be some reason to suppose that intuitions on twin-earth cases will correlate to a reasonable degree with the semantic facts. If the reference of a word is in any part determined by the dispositions of speakers to apply that word under various circumstances, then we should expect twin-earth intuitions to be a
fair guide to semantic stability - if our intuitions are a reasonably reliable guide to those dispositions. It is an open empirical question, however, whether they are.

Even setting this worry aside, it is not obvious that this argument for the stability of philosophical terms succeeds. For one thing, the account of categorial understanding is in some respects puzzling. Take the example of water. If Bealer’s account is correct, then we are able to determine, *a priori*, the general category to which water belongs – we know *a priori* that water is a compositional kind. But why should we expect to have *a priori* access to such a fact? Bealer first suggests that a solution would follow were we to accept the following, simplified account: the fact that ‘water’ refers to a compositional kind was determined by the intentions of some group of initial water-baptizers, when introducing a term to refer to the watery samples of their acquaintance, to pick out a compositional kind. “On this oversimplified account, the baptizer would then be in a position to know *a priori* that, if it exists, water is a purely compositional stuff,” (Bealer 2002, 108). Though this would indeed explain the baptizers’ ability to understand the categorical content of their term *a priori*, one begins to wonder at the fact that the intentions of these original water-baptizers seem to have meshed so well with the understanding of watery samples provided by modern chemistry.

More plausibly, the term ‘water’ was introduced before anyone knew whether the samples of water they had encountered had any compositional structure in common. It seems unlikely, then, that ‘water’ was consciously introduced as a term picking out a compositional stuff. Bealer notes this, and proceeds to attempt a correction of the oversimplifications present in the initial suggestion. Bealer suggests that, instead, ‘water’
was introduced by means of a categorial concept composed of an ordered list of conditionals –

“for example, the concept of being the stuff S such that (1) if all and only samples of S in our acquaintance are samples of some purely compositional stuff, then S is that compositional stuff; (2) if there is no such purely compositional stuff and if, instead, all and only instances of S in our acquaintance are instances of some not too complicated impure compositional stuff (akin to the impure compositional stuff jade), then S is that impure compositional stuff; (3) if there is no such pure or impure compositional stuff and if, instead, all and only instances of S in our acquaintance are instances of some not unwieldy macroscopic stuff, then S is that macroscopic stuff; and so forth” (Bealer 2002, 109).

Importantly, if this account is correct, then whether a natural kind term picks out a compositional or a macroscopic kind is dependent on contingent features of the actual world.

One consequence of this modified account is that we should expect to find cases where, had the science turned out differently, we might have been inclined to conclude that a certain natural kind term picked out a stuff belonging to a different category than what the term actually picks out. Specifically, we should expect to find intuitions of the following form: for some time t before the occurrence of any scientific discoveries relevant to the stuff S picked out by expression e, 52 there could have been a community of speakers in an epistemic situation qualitatively identical to ours at time t who use e to pick out a stuff, S’, belonging to a different category than S.

‘Jade’ provides a suitable example. In the actual world, ‘jade’ is a term which picks out a disjunctive kind – x is an instance of jade iff x is an instance of jadeite or x is an instance of nephrite. Prior to the 19th century, however, it was not known that the stones being referred to as ‘jade’ in fact comprised two distinct mineral types. Imagine a community of speakers in the actual world just prior to this discovery – say, in Britain in

52 Here, of course, expressions must not be understood as being individuated by their semantic content, but rather by their phonological/orthographic properties, or something similar.
the 18th century. Imagine a possible world containing a group of speakers in an epistemic situation qualitatively identical to that of the 18th century Britons, but in which the jade-ish samples of the speakers’ acquaintance are all composed of jadeite. In fact, there is no nephrite at all on their planet – only jadeite. There is, however, nephrite on a small, uninhabited planet at the furthest reach of the galaxy.

According to Bealer, we should have fairly strong intuitions about whether, in the mouths of members of the described nephrite-deprived community, the term ‘jade’ simply picks out jadeite, or whether it picks out the disjunction of jadeite and nephrite. If ‘jade’ has a categorial conditional ranking according to which the compositional is to be given priority over the macroscopic, then ‘jade’, as used in the nephrite-poor world, refers only to jadeite – just as ‘water’ refers only to H2O on our world, even if we subsequently discover some watery stuff with the composition XYZ. The watery samples present in our environment when the term ‘water’ was introduced were all samples of H2O, and it is the common compositional features of those samples which determined the extension of the term. So too for ‘jade’ on the nephrite-poor world. On the other hand, if the term ‘jade’ is such that the macroscopic has priority over the compositional, then ‘jade’ refers to jadeite and nephrite at the nephrite-poor world, regardless of whether any nephrite appeared in the original samples. And, if Bealer’s account of our a priori understanding of these categorial conditionals is correct, then which of these alternatives obtains should be quite intuitive. Unfortunately, I at least find myself lacking clear intuitions on the case.

Even putting the unclarity of such intuitions aside, Bealer’s ordered conditional proposal seems hardly more plausible than the original account. Our predecessors did not
know *a priori* that water would turn out to be a compositional kind; that knowledge required empirical investigation. Why, then, should we expect that a group of ancient water-baptizers should have *a priori* access to some ordered list of conditionals of the sort Bealer suggests? How might those ancients have correctly intuited that the compositional carving-of-nature should be given priority over the macroscopic? It is, after all, plausibly a contingent empirical fact that compositional kinds play an important role in scientific explanation; we might have inhabited a world where the kinds that entered into nomologically robust generalizations were all macroscopically individuated.

A more probable hypothesis, to my mind, is that our understanding of the categorial content of our concepts reflects a good deal of empirical expertise with the relevant kinds; as our scientific knowledge increases, we discover whether the samples of our acquaintance best characterize a compositional kind, a functional kind, or some other sort of kind. In some cases – as with ‘jade’ – historical, cultural, and social factors may influence our judgments regarding the term’s semantic properties, as well. Had the disjunction of jadeite and nephrite enjoyed less commercial importance in 19th century Europe, we might have adopted separate terms for the two substances once they were distinguished.

Even if Bealer’s ‘ordered conditional’ account turns out to be correct, however, it clearly represents a significant weakening of the previous claim that we have *a priori* knowledge of the fact, e.g., that water is a compositional kind. The categorial content of semantically unstable concepts is, at least to some degree, dependent on contingent facts about our environment. To the extent to which the categorial content of these concepts are dependent on contingent facts about the environment, that categorial content is,
strictly speaking, partially *a posteriori*. Though we may, on Bealer’s modified account, have *a priori* access to a set of ‘priorities’ – e.g., should there be both a compositional and a functional kind in the vicinity of the relevant samples, it is *a priori* that the compositional kind is the true referent – the actual categories to which even natural kind concepts belong remain unknowable until provided by empirical investigation.

Worse, though, the ordered conditional account suggests that some terms that *appear* to possess semantic stability might in fact have ordered conditional structure, as well. Certain categorial contents, notably the macroscopic and the functional, appear to lend semantic stability to the terms possessing them. For instance, if ‘watery stuff’ picks out a certain macroscopic kind – the kind characterized by wetness, clearness, drinkability, etc. – then it appears to refer to what it does independently of the contingent features of the actual world. Twin earth doppelgangers refer to the same sorts of things that we do when using the expression ‘watery stuff’, even though the watery stuff of their acquaintance is XYZ – the expression is semantically stable. Similarly for, e.g., ‘table’. But if Bealer’s ordered conditional account is correct, the semantically *un*stable term ‘water’ might have referred to a macroscopic kind, had the world been such as to force us far enough down our list of ordered conditionals. Had there been no unified compositional kind or ‘jadish’ kind in the vicinity, ‘water’ would have picked out watery stuff. And presumably, it would have appeared (at least prior to reflection) to be semantically stable. Only on further consideration would it have become clear that the appearance of semantically stable behavior was contingent on a lack of compositional unity among the watery samples of our acquaintance at the time of the introduction of the
term. This ‘further consideration’ might well include consideration of a thought experiment quite like that given for jade above.

Recall that, if a term T is semantically unstable, then we should intuit that it is possible for there to be two groups of speakers in epistemically identical situations which nonetheless pick out different things with T. Insofar as T possesses an ordered conditional structure to its categorial content, we should be able to intuit that it is possible for two such groups to pick out different kinds of stuff with T. For reasons I will give below, I am inclined to believe that intuitions of this form can arise for many philosophical concepts. If this is correct, then it would appear that Bealer ought to conclude that these philosophical terms share the ‘ordered categorial conditionals’ feature that he assigns to semantically unstable terms – after all, for Bealer, intuition provides the primary data for determining a concept’s categorial content. Since an ‘ordered conditional’ categorial content, by its very nature, introduces a degree of a posteriority to the understanding of the reference of such terms, it follows that at least some philosophical concepts are unstable. I will now attempt to provide some ‘twin earth’ scenarios for philosophical concepts that, to my mind, are more-or-less intuitive.

As a preliminary, it should be emphasized that the community of speakers under consideration should not be required to inhabit a twin earth within the actual world; that is, we should be permitted to appeal to cases involving non-actual possible worlds. The reason is this: semantically stable terms are meant to underwrite the autonomy of philosophy. Now, a term is semantically stable if and only if the external environment makes no contribution to its meaning. Such a definition only underwrites autonomy if ‘external environment’ is read as encompassing facts about, e.g., the laws of nature.
Facts about the laws of nature are not available *a priori*; should the meaning of a given term depend on such facts, we will not be able to investigate the kind designated by that term without input from the sciences. We should, then, be permitted to appeal to cases in which laws of nature differ between the two communities (unbeknownst to them, of course).

There are, I think, a number of plausible cases of this sort for philosophical terms. After all, there are a great many philosophical theories whose popularity has diminished as a result of developments within the sciences. For at least some of these, we may be inclined to say that the theories would have been correct, had the science turned out differently. As an example, consider a Cartesian substance dualism about consciousness. The majority view among philosophers is that substance dualism is false. One good reason for rejecting substance dualism is the widely-held view that the physical world is causally closed – i.e., the view that every physical event has a physical cause. Insofar as this is read as a claim about the sorts of entities currently recognized by physics\(^53\), it is surely both contingent and empirical. Further empirical investigation might reveal that the ontology of current physics is incomplete.

Suppose that we do, in fact, inhabit a world where the closure principle just specified holds. Nevertheless, it might have turned out that the closure principle was false. Imagine a possible world, \(w\), in which the physical world is not causally closed and in which there exists a non-physical mental substance in which consciousness inheres. Imagine there is a community of speakers on \(w\) in a situation epistemically

\(^{53}\) I.e., if we set aside interpretations of physicalism formulated in terms of future, completed physics – under such interpretations, physicalism is plausibly trivially true.
identical to ours\textsuperscript{54}, and that all members of this community possess immaterial minds composed of the aforementioned mental substance. Is it absurd to suppose that, for the inhabitants of \( w \), ‘consciousness’ means something different than it does for us – that it picks out a non-physical mental state, rather than a physical state? If not, then perhaps ‘consciousness’ is not semantically stable after all.

My own intuitions on this case are muddy. Perhaps ‘consciousness’, in both the actual world and world \( w \), applies to states in both worlds. Perhaps all conscious states in the actual world are physical, but non-physical states in other worlds would fall under the term ‘consciousness’, as well. At best, however, this would seem to indicate that ‘consciousness’ parallels ‘time’ – that there is something like a ‘generic’ use of the term that is semantically stable. For debates over the mind-body problem, however, the non-generic use would clearly be the most relevant. Why else would philosophers be inclined to argue that substance dualism is false? If it is indeed the non-generic use that interests philosophers, then it is implausible that autonomy holds for philosophical theories of consciousness.

A quite similar case can be constructed for the moral term ‘good’. In the actual world, under the assumption that physicalism is true, the property of goodness must be some sort of physical property. Let’s imagine it is in fact identical to the property of causing more pleasure than pain. One can nevertheless imagine a “Moorean” possible world in which physicalism is false, and in which a community of speakers in an identical epistemic situation to our own uses the term ‘good’ to refer to a simple, indefinable, non-natural property supervening on exactly those states which cause more pleasure than

\textsuperscript{54} In order to guarantee epistemic similarity, one might imagine the two communities at a time \( t \) before the development of modern physics.
pain. The same considerations as above hold – if one is inclined to claim that the properties in both worlds fall under our term ‘good’, then this at best shows that there is a generic use of ‘good’ whose relevance to ethical debate is likely minimal. Else, debates over moral naturalism would be beside the point.

At the very least, I take these examples to show that the claim that the majority of philosophical concepts are semantically stable is open to serious debate. I suspect that a more creative soul – and one with a deeper understanding of the history of the interactions between science and philosophy – could greatly multiply the examples. If so, then Bealer’s claims to autonomy and authority will begin to look deeply suspect.

If a given philosophical concept is unstable, then at best we can achieve *a priori* determinate understanding of that concept in the limited sense relevant to natural kinds – i.e., a determinate understanding of its categorial content. As noted above, however, this may merely amount to knowledge of some conjunction of conditionals. Alternately, if my suspicions regarding e.g. the contingent cultural factors influencing the reference of ’jade’ are indicative, we may not be able to pursue *a priori* determinate understanding in any sense – it may be that any substantive understanding of the categorial content of the concept requires empirical investigation. In either case, the prospects for a fully autonomous and authoritative philosophical theory of the concept at hand would be dim.

Though these considerations are not, of course, decisive, they do suggest that Bealer’s account of concept possession is insufficiently motivated when considered as an independent argument for the autonomy and authority of philosophy. Had Bealer’s Argument from Evidence proved successful, this might have provided grounds for a claim that some account must be given of the strong modal tie underwriting intuition’s
status as a basic source of evidence. From there, it might have been possible to claim that an account of concept possession resembling the one given in the argument from concepts must be, at the very least, broadly correct. However, as I hope to have shown in the first half of the paper, the Argument from Evidence can be questioned at numerous points. We should therefore feel no pressure to accept Bealer’s account of concept possession, nor to endorse the theses of Autonomy and Authority.

It is worth noting that no argument has been made that there is no a priori component to our understanding of natural kinds. In his (1987), Bealer claims that a proponent of scientific essentialism must advert to intuition to defend the claim that natural kind identities are necessary if true. But the a priority of this fact (if it is indeed a priori) doesn’t imply the sort of autonomy or authority that Bealer is ultimately concerned with. The indispensability of intuition for the defense of scientific essentialism does not entail that philosophical questions about categories like ‘belief’, ‘knowledge’, or ‘goodness’ can be answered without the aid of empirical investigation. Instead, it suggests the following: just as we need science in order to do our philosophy, we need philosophy in order to do our science. If anything, such a view further blurs the distinction between philosophical and scientific methodologies.

I am, myself, very sympathetic to the position that the boundary between philosophy and science is extraordinarily fuzzy. Such a view, however, is in serious tension with Bealer’s characterization of philosophy as an autonomous, authoritative discipline. As I have attempted to show, Bealer’s arguments from Evidence and from Concepts fail to motivate the theses of autonomy and authority. Worse, though, even certain background assumptions to Bealer’s project strike me as misguided. Bealer
attempts to characterize the questions of philosophy as being wholly independent from questions asked by the sciences; in several papers, he notes that philosophical questions possess certain presumably definitive characteristics, including universality, generality, and necessity. However, it’s quite unclear to me that these features are unique to philosophical inquiry.

Bealer claims that philosophical questions are universal, in the sense that any philosopher, regardless of cultural or historical context, would find them of significant interest, at least once introduced to the required concepts. This is a deeply empirical claim, and I suspect a false one; but even assuming its truth, it’s unclear why a similar feature should not hold for scientific questions. Bealer claims that philosophical questions are general, in the sense that they are phrased in general terms and do not advert to particular individuals or entities. But such a feature is quite clearly true of many scientific questions, as well – scientists frequently aim to characterize regularities and generalities, such as the solubility of various mineral kinds or the behavior of primates under stress. Finally, Bealer claims that philosophical questions are necessary, in the sense that the answers to these questions hold of necessity, rather than contingently. This feature, at least, is absent in a large number of scientific cases – it is not necessary (in the metaphysical sense) that water dissolves salt, for example. But in many other cases, scientific investigations are concerned with uncovering necessities. It is a necessity, for example, that water is composed of H$_2$O. Mutatis mutandis for the essential characteristics of other natural kinds.

It is my belief that any attempts to demarcate a privileged methodology or subject matter for philosophy are doomed to failure. Rather than bemoan the fate of the theses of
autonomy and authority, however, we philosophers may take pride in the extent to which our discipline so essentially connects to others. It is this very interconnectedness, after all, that makes philosophical investigation so exciting, so challenging, and so rewarding.
Chapter 5

Truth and Theoretical Utility

There are a number of related arguments, which I call ‘constitutivity’ arguments, that aim to defend the use of intuition in philosophy via a claim that the reliability of intuition is guaranteed by its very nature. Typically, this claim is cashed out in terms of the existence of some sort of constitutive relationship between intuitions and meanings or between intuitions and concepts. Alvin Goldman, for instance, writes that “It’s part of the nature of concepts… that possessing a concept tends to give rise to beliefs and intuitions that accord with the contents of the concept” (Goldman 2007, 15). Similarly, George Bealer writes that “it’s necessary that the comprehensive theoretical systematization of a subject’s intuitions in [suitably high-quality cognitive conditions] is largely true” (Bealer 1996, 130-131).

The aim of this chapter will be to examine the prospects for arguments of this sort. Rather than focusing on the plausibility of the accounts of meanings and of concepts that these accounts require, I will assume their truth for the sake of argument and focus on their consequences. It will be my contention that, even if we grant the existence of constitutive ties between intuition and meanings or concepts, traditional intuition-based methodology will not automatically be thereby vindicated. Though constitutive ties, if they exist, may guarantee the overall reliability of intuition, they do not necessarily address the question of whether the categories described by our intuitive classificatory practices are appropriate ones for our theoretical purposes. The constitutive tie between intuitions and meanings (if it exists) might, for instance, provide us only with
a reliable ability to delineate the boundaries of categories that play no important explanatory role in our philosophical or scientific theorizing.

This chapter will focus primarily on the account of the indispensability of conceptual analysis given by Frank Jackson in his book *From Metaphysics to Ethics*. As will be seen, Jackson’s account of the relation between intuitions and meanings invokes a broadly descriptivist account of meaning – one which guarantees that our intuitions about the meanings of our terms will be by-and-large true. However, this by itself leaves open the possibility that our terms will turn out to refer to, e.g., wildly disjunctive, gerrymandered, or otherwise theoretically non-optimal categories. The degree to which this is a problem for Jackson’s account will depend upon one’s views on the aims of philosophy. If one sides with philosophers who (to adopt terminology introduced by Alvin Goldman and Joel Pust) take a broadly ‘extra-mentalist’ approach to the aims of philosophy, claiming that the targets of philosophical theorizing are independent of our psychology, then Jackson’s account potentially becomes deeply problematic (or so I will argue). If one takes a ‘mentalist’ approach, claiming, as Goldman himself does, that the targets of philosophical theorizing are psychological, then the problem may be averted. However, I will argue that the mentalist approach at best only characterizes one facet of philosophical investigation.

1. Classical descriptivism and the reliability of intuition

There is something very attractive about the idea that intuitions reflect an *a priori* understanding of the circumstances under which a concept applies to a described object
or scenario. As we saw in the last chapter, George Bealer attempts to argue that intuitions are reliable because we possess *determinate* understanding of our concepts – that is to say, for a given object *x* and determinately understood concept *F*, presentation of a suitable description of object *x* will allow us to accurately intuit whether or not it falls under concept *F*. Unfortunately, the actual mechanism behind this proposed determinate understanding remains fairly opaque; just why should it be that we would possess such reliable knowledge of the boundaries of *F*-hood? Bealer offers little by way of argument as to how or why our actual philosophical concepts are in fact possessed determinately, or at least possessed determinately enough to allow us to conclude that philosophy may rely on intuitions as evidence.

In fact, Bealer’s approach attempts to run the other way around; that intuitions are evidence is taken to be unassailable, and Bealer’s confidence in determinate concept possession seems to emerge via something like abduction. If we possessed concepts determinately (enough), this would explain intuition’s evidential status. Therefore, it must be the case that we possess concepts determinately (enough). Of course, those who remain unconvinced by Bealer’s claim that intuitions are indispensable as a source of evidence will not be satisfied by this mode of argument.

Can we supplement Bealer’s argument in a way that would motivate, rather than assume, the presence of a strong modal tie between our intuitions and truth? Let’s return to the example, given in the previous chapter, of the woman who determinately possesses a concept ‘multigon’. Recall that the woman introduces, through use, a term ‘multigon’; she applies it to various closed figures but has not, as of yet, applied or withheld it from triangles or rectangles. Bealer’s claim is that if the property of being a multigon is
identical to the property of being a polygon with five or more sides, then, intuitively, the woman will judge that a triangle is not a multigon. What could possibly explain her ability to do so? Here is one suggestion: determinately possessing the concept ‘multigon’ involves possessing a mental description ‘polygon with five or more sides’; further, the reference of ‘multigon’ is determined by this description. The woman’s possession of this mental description leads her (in suitably high-quality cognitive conditions) to judge that x is a multigon just in case x is a polygon with five or more sides; and, because this description determines the reference of ‘multigon’, x is a multigon just in case x is a polygon with five or more sides. So the woman judges that x is a multigon iff it is true that x is a multigon.

Classical descriptivism, then, seems to be the obvious way to elucidate Bealer’s hypothesized strong modal tie. Even better, descriptivism would straightforwardly allow philosophical terms to be ‘semantically stable’ in the way Bealer’s account requires – Bealer could simply argue that philosophical terms are associated with descriptions that pick out the same extensions regardless of contingencies of the speaker’s external environment. I am not here suggesting that Bealer is, in fact, a descriptivist, either explicitly or implicitly. But classical descriptivism is the most obvious route to the necessary reliability that constitutive approaches like Bealer’s require.

It is clear that classical descriptivism would imply that many of our intuitions are generally reliable, at least in suitably good cognitive conditions. According to classical descriptivism, speakers associate natural language expressions with descriptions, and these descriptions determine the reference of the expressions. Imagine a speaker S associates with an expression φ a description α. The extension of φ, according to
descriptivism, is the set of entities that fall under the description $\alpha$. Now, if $S$ is presented with a thought experiment in which it is stipulated that a certain object has such-and-so features, and if in virtue of those features $S$ can determine that the object falls under description $\alpha$, and if $S$ were to be asked if that object counts as a $\varphi$, it seems that so long as cognitive conditions are suitably good $S$ will consult her associated mental description and judge that the object counts as a $\varphi$. And, descriptivism entails that the object is a $\varphi$; so $S$ judges truly. And, of course, if $S$ is presented with a thought experiment in which an object is stipulated to be such as to not fall under $\alpha$, $S$ will judge that it is not a $\varphi$ – and she will be right. So long as the thought experiment provides her with enough information to determine whether or not the conditions laid down by $\alpha$ are met, $S$’s identification intuitions will be reliable.\textsuperscript{55}

It is, of course, this very feature of classical descriptivism that underlies the success of Kripke’s argument from error. Classical descriptivism should make significant error impossible; if a word’s meaning were determined by the mental description of the speaker, then a speaker should automatically be accurate in her beliefs about the referents of her terms insofar as those beliefs derive from the mental description which determines reference. For a great many terms – famously, names and natural kind terms – this is quite clearly not the case. My utterance of the word ‘elm’ refers to elms, even if the mental description I associate with ‘elm’ in fact uniquely describes beeches.

Many appear to side with Bealer in taking Kripke’s arguments to apply primarily to names and natural kind terms, while failing to apply to most philosophical terms.

\textsuperscript{55} There is a complication here; in order for the concept under consideration to be fully definite, the description $\alpha$ must be such that every possible case either falls under it or fails to fall under it. Let’s simply grant for the sake of argument that the descriptions speakers associate with natural language expressions do in fact have this feature.
Nonetheless, contemporary philosophers with descriptivist leanings are still under some pressure to identify a mechanism for ensuring some level of conceptual fallibility – infallibilism about intuition has fallen quite out of fashion. One solution is a ‘cluster’ theory, according to which the referent of a term need satisfy only a majority of the descriptive information associated with that term. This strategy would lessen the reliability of intuition somewhat, though a good deal of constitutively-grounded reliability would remain. In the next section, I will examine the most sophisticated contemporary attempt to ensure the necessary reliability of intuition via a broadly descriptivist strategy – that of Frank Jackson. I will then attempt to show that, depending on one’s views on the aims of philosophical theorizing, this necessarily reliability may offer little comfort to those seeking to legitimize the current role of intuition in philosophy.

2. The two-dimensional approach to constitutivity

We will begin by examining the basics of Jackson’s two-dimensional account of conceptual analysis, as presented in his *From Metaphysics to Ethics*. As a preview, Jackson claims that the two-dimensional framework highlights an aspect of meaning which is independent of any knowledge of which world is actual, and which is accessible via folk reactions to descriptions of possible scenarios. This aspect of meaning provides, for each expression, a function mapping microphysical descriptions of possible worlds to extensions within those worlds.
The root of Jackson’s theory is the assertion that philosophy requires ‘serious metaphysics’. Serious metaphysicians work with a limited ontology, countenancing only the most fundamental of ingredients. Most terms, and practically all philosophically interesting terms, will not be explicitly mentioned in this ontology. However, Jackson claims, the existence of ‘higher-level’ entities may be entailed by the basic ontology, in the sense that facts about the higher-level entities may be supervenient on facts about the lower-level entities.

Take belief as an example - in order for beliefs to be supervenient on basic physicalist ontology in the relevant way, it must be the case that fixing the (micro-) physical way the world is thereby fixes the belief-y way the world is. If beliefs are supervenient on physical facts in this way, then this justifies beliefs’ inclusion in the serious metaphysician’s implied ontology. More generally, higher level entities are granted ‘entry by entailment’ – if the existence of a given entity is entailed by facts specified in the terms of the fundamental ontology, then it is granted a legitimate place in the implied ontology of the serious metaphysician. For each entity described in a high-level vocabulary, the serious metaphysician must either ‘locate’ that entity by showing how its existence is entailed by facts described in the vocabulary of her basic ontology, or she must eliminate the higher-level entity from her ontology entirely.

It is here that conceptual analysis enters the picture. According to Jackson, “conceptual analysis is the very business of addressing when and whether a story told in one vocabulary is made true by one told in some allegedly more fundamental vocabulary” (Jackson 1998, p. 28). It is only once we have used conceptual analysis to locate our subject that will we be able to answer philosophical questions about it. In
order to determine whether all beliefs are brain states, for instance, we first need to know what counts as a belief – and in order to find out what counts as a belief, we must analyze the word ‘belief’ to discover its connection to our fundamental vocabulary.

Jackson’s version of conceptual analysis, as is traditional, centers on intuitions about possible cases. The analysis of a term is to be extracted from a set of data comprised of folk intuitions about the application of a term in a given scenario. Also, again as traditional, the intuitions and the analysis derived from them are viewed as fully a priori. Where Jackson’s story departs from the classical version is in the details - it addresses Kripke’s worries by making use of mechanisms for rigidification, and makes use of the framework of two-dimensional semantics to lend plausibility to the claim that we can achieve analyses of our terms entirely without use of empirical knowledge.

As noted in the previous section, classical descriptivism faces the difficulty of explaining how we can successfully employ names and natural kind terms despite knowing very little about their referents. One way to escape this problem while maintaining a descriptivist framework is to hold that the mental descriptions associated with names and natural kind terms incorporate some manner of rigidifying feature. Take ‘water’ as an example. Jackson suggests that the mental description associated with ‘water’ is something like ‘water is the watery stuff of our acquaintance’ (Jackson 1998, 52) – in other words, water is the actual watery stuff around here. ‘Watery stuff’ refers, in every world, simply to any stuff at that world that is wet, clear, drinkable, falls from the sky, and so on – thus, it could refer to H₂O in this world, but XYZ in another.

However, addition of reference to the actual world in the description associated with
'water' prevents such reference-shifts from occurring – ‘water’ thus picks out H₂O in all worlds.

As Jackson reads Kripke, it is acquaintance with the local watery exemplars during a ‘baptizing’ event that determines how the referent of the rigidly designating term ‘water’ is fixed. Thus, had we occupied a world where the watery stuff had the chemical structure XYZ, the baptizing would have fixed the referent of ‘water’ as that kind of stuff – and water would have been, necessarily, XYZ. This is due to the fact that, had we lived in an XYZ-world, the stuff fitting our mental description – the ‘actual’ watery stuff ‘around here’ – would have been XYZ. Thus, building in reference to a feature of the actual world in this way guarantees that ‘water’ will act as a rigid designator, just as Kripke’s thought experiments suggest it should.

The incorporation of the ‘actuality’ operator into the mental descriptions associated with our terms suggests a ‘two-dimensional’ semantics. Two-dimensional semantics holds that the answer to the question of what a term refers to at a given world is two-fold. We can consider what the term refers to, given that the world in question is counter-factual – this will give us the C-extension of the term, the ‘counterfactual extension’. It is this approach to the reference question that underlies statements of a posteriori necessity like “Water is H₂O in every possible world”. However, we can also consider what a term refers to, when a target world is considered as actual. This results in the A-extension, or ‘actual extension’. Upon this approach, ‘water’ does not refer to H₂O in every possible world. Instead, ‘water’ refers to whatever natural kind underlies the exemplars of ‘watery stuff’ in the target world.
To the notions of C-extension and A-extension, we can add corresponding notions of C-intension and A-intension. The C-intension of a word is a function that takes possible worlds and assigns to each a C-extension; the A-intension takes possible worlds and assigns to each an A-extension. The C-intension of ‘water’ takes each world and picks out the instances, in that world, of the referent of ‘water’ at the actual world – i.e., for each world, it picks out the instances of H₂O. Thus, to know the C-intension of water, one must know that ‘water’ refers to H₂O in the actual world. ‘Water’ refers to H₂O in the actual world because the watery stuff the baptizers for ‘water’ were acquainted with was H₂O – thus, to know the C-intension of ‘water’, one must know that the watery stuff in the actual world is H₂O.

A-intensions, however, are not tied to information about the actual world in this way. Here, then, is Jackson’s crucial argument for the tractability of a priori conceptual analysis. Since the A-intension of a term picks out A-extensions at worlds without regard to any characteristics of the actual world, knowledge of A-intensions may be obtained merely by understanding the term at hand – i.e., may be obtained strictly a priori. Jackson claims that conceptual analysis, by examining folk intuitions about extensions of terms in possible scenarios, reveals the implicit understanding of the folk and uncovers the A-intensions of their terms. In this way, conceptual analysis reveals a priori truths about the meanings of our words.

Jackson’s view of conceptual analysis is clearly a constitutive one. Speakers possess descriptive conceptual information which enables them to intuit category membership when provided with a rich enough specification of a hypothetical scenario. This same descriptive information determines the A-intensions of their terms. These A-
intensions, in turn, determine the A-extensions of terms. Thus, a subject presented with a sufficiently detailed thought experiment will be able to accurately pick out members of the A-extensions of her terms. Of course, when a term’s A-extension does not match its C-extension (as in the case of ‘water’), the subject may not be able to accurately pick out members of the C-extension of that term at a world. However, there will still be a subset of beliefs for which her intuitions will be a reliable guide to truth. This subset is the set of what Chalmers and Jackson (2001) call ‘application conditionals’ – conditionals of the form E -> C, where C is a statement characterizing the extension of a given concept, and E is sufficiently detailed information about a given possible world.

The application conditional is essentially a restatement of the a priori knowledge we have of the A-intensions of our terms – we know that if some situation E is actual, then the extension of term T is such and so. If the situation in which the local watery stuff is XYZ is actual, then the extension of ‘water’ is XYZ. Note the similarity to the general method of the thought experiment – one might rephrase Gettier’s contribution to the analysis of knowledge as G -> ~K, where G is a description of a Gettier scenario and K is a claim that a certain mental state of that scenario’s protagonist falls under the extension of ‘knowledge’.

Jackson’s two-dimensionalism provides a means for accommodating Kripkean intuitions by incorporating an aspect of meaning in which descriptive conceptual information is rigidified. Meanwhile, it retains an a priori link between mental descriptions and meanings in the form of A-intensions, which can be expressed through application conditionals. However, as I will argue in the next section, it leaves a crucial question unanswered.
3. The question of theoretical utility

The following is a feature of both classical descriptivism and Jackson’s two-dimensionalism: no constraints are placed on the naturalness or theoretical importance of the extensions (or A-extensions) of our terms. Imagine that I introduce by stipulation a term, ‘lree’. The mental description I associate with ‘lree’ is as follows: x is a lree iff x is a streetlamp or a tree. Whenever I am presented with a sufficiently detailed thought experiment, I’m able to consult my intuitions and judge whether the entity specified is a lree or not – I judge ‘x is a lree’ when and only when the thought experiment either specifies that x is a streetlamp or specifies that x is a tree. According to classical descriptivism, ‘lree’ refers to anything satisfying the description ‘streetlamp or tree’ (for this is the mental description I associate with the term), and so ‘x is a lree’ is true when and only when x is a streetlamp or a tree. Mutatis mutandis for Jackson’s account, so long as the A-extension of ‘lree’ matches its C-extension (and there is no obvious reason why it should not). My intuitions, then, will turn out to be a very reliable guide to true beliefs about lrees.

True beliefs about lrees, however, don’t really seem to be particularly interesting or important; they don’t seem to have a role to play in any scientific or philosophical theory. Why not? There are several possibilities – the notion of a lree isn’t predictive or explanatory, perhaps, or not simple enough, or otherwise not ‘natural’ enough. Or maybe it’s simply that there are other notions nearby – the notion of a streetlight, and the notion of a tree – that would do better in some sense. The fundamental observation is simply
that there are some categories to which accurate, explanatory theories do not advert –
disjunctive categories like lree, empty categories like phlogiston and élan vital,
explanatorily impotent categories like brachycephalic (a category of skull shape thought
by practitioners of craniometry to indicate low intelligence and moral inferiority), etc.
There are some ways of carving up the world which are defective, and which do not
facilitate successful theorizing.

The idea that our epistemological goals are not satisfied merely by possession of a
high proportion of true beliefs is not a novel one; reliability is clearly not the be all and
end all of cognitive success. Goldman (1986), for instance, notes that one might have a
very reliable process that nonetheless produces very few beliefs overall, or that operates
extremely slowly, or that produces only beliefs that are unrelated to the one’s needs or
interests. A system that produces nothing but highly accurate reports on the growth rate
of grass on a nearby lawn is not particularly worthy of epistemic praise. On a
descriptivist account, true intuitions come cheap; it does not follow that descriptivism
entails that those intuitions always lead to epistemological success.56

I wish to digress for a moment to address the question of whether the categories
which underwrite theoretically useful beliefs are just those categories which are natural
kinds. There is a fair amount of disagreement over just what it takes to be a natural kind;

56 Stich (1996), Bishop and Stich (1998), and Mallon, Machery, Nichols and Stich (forthcoming) have
argued against drawing substantive conclusions about ontology from one’s account of reference. Stich
(1996), for instance, argues that it is far from obvious that there is a single privileged word-world relation
upon which existence claims can be founded, and that we should therefore resist drawing ontological
conclusions from semantic premises. My argument is somewhat different from these. I am, in fact,
relatively comfortable with the idea that there might be a privileged word-world relation, and if there is
such I find it plausible that the truth of existence claims could be argued for on the basis of that relation.
My claim in this paper is that, while the truth of statements like ‘beliefs exist’ may well follow fairly
straightforwardly from a given account of reference (plus relevant empirical premises), the epistemological
worth of such statements does not. The upshot, however, is similar – don’t try to read your metaphysics off
of your semantics.
However, on some conceptions – particularly ones which take natural kinds to be in some sense ‘mind-independent’ – the set of natural kinds is much smaller than the set of categories which I have in mind. There are, for instance, some categories which are arguably not natural kinds in the ‘mind-independent’ sense but which may play important roles in explanation of one sort or another – marriage, religion, and currency, to name a few. The disciplines of sociology, economics, etc., are genuinely explanatory even if the kinds they invoke in their explanations are counted as non-natural on the mind-independence criterion. And presumably, true philosophical theories will be genuinely explanatory even if it should turn out that the kinds they invoke are mind-dependent in a similar sense.

Alternatively, taking a cue from Quine (1969), we might speculate that the natural kind categories are just those expressed by projectible predicates – predicates which can occur in good inductive inferences. The notion of projection is a thorny one, however, and though it likely allows in at least some of the explanatory yet ‘mind-dependent’ categories, it seems to me that the criterion may be too broad for our purposes. Take, for instance, the ‘emeroundsquare’ – something is an emeroundsquare iff it is an emerald or a round square. All the emeroundsquares I have observed to date have been green, and I’ll do quite well if I infer that all emeroundsquares are green. Yet the term ‘emeroundsquare’ is in some sense not appropriate for the purposes of scientific or philosophical theorizing. Further, we might note that explanatorily impotent categories like brachycephalic often enter into at least some valid inductions - inductions regarding which hats will fit, perhaps.

\[57\] Thanks to Richard Dub for help formulating this example.
With these considerations in mind, I prefer to avoid characterizing the set of categories relevant to theoretical success in terms of the notion of a natural kind. At the same time, of course, the prototypical natural kinds are prime examples of the sorts of categories successful theories will invoke. Let us call the categories which are useful in scientific or philosophical theorizing ‘explanatory kinds’, and simply note that the set of explanatory kinds at the very least probably includes at least the set of prototypical natural kinds as a subset.  

A few further observations about explanatory kinds will be helpful in what is to come. First, a category can be explanatory with regard to one area of inquiry while failing to be explanatory with regard to another. Race, for instance, is arguably not an explanatory notion in biology, but may be explanatory with regard to sociology (I am thinking here of, e.g., the study of racism). Second, a category can be explanatory locally but not globally; C-fiber firing can explain pain in humans without explaining pain simpliciter. Finally, explanatoriness is a matter of degree – neither ‘reliably formed true belief’ nor ‘true belief formed in accordance with one’s evidence’ characterize kinds which are as defective as ‘lree’, yet one of these two kinds may be a better candidate for a role in our epistemological theorizing, and hence be more explanatory than the other.

58 The notion of explanation itself will also be left vague. This, however, is because the general worry I am trying to characterize for constitutive approaches to intuition should arise on any plausible account of explanation. It also, by the way, arises for other values of interest – if some philosopher happened to be concerned with aesthetic appeal rather than explanation, she could restate the general worry as follows: constitutive approaches to intuition may guarantee reliability, but they cannot guarantee that our concepts carve the world in the most aesthetically appealing way.

59 Ultimately, there need not be any particularly natural or simple way of specifying the sorts of categories that feature in successful theories – indeed, they need not have any deep feature in common other than occurring in successful theories. For current purposes, it will be sufficient to note that there exist some categories, like lrees, that very plausibly do not have any sort of theoretical utility, and that descriptivism does not guarantee that our intuitive classifications will not be similarly defective.
To return at last to the main thread of argument: descriptivism allows beliefs which are true but which may involve only uninteresting, non-useful categories, and which may therefore be epistemologically sub-par. But, one might argue, descriptivism also allows beliefs which do invoke interesting categories – so what’s the problem? The problem is that descriptivism does not, by itself, guarantee that any of the intuitions I have regarding philosophical terms like ‘knowledge’, ‘consciousness’, ‘person’, and so on, invoke interesting categories. How do I know that the description I associate with ‘knowledge’ picks out a sufficiently explanatory kind? Note that the problem is not just that some person might have a deviant usage for the term ‘knowledge’ – i.e., that someone might associate a description like ‘false belief’ with the term knowledge. Presumably we need nothing more than deference to experts to defuse that particular worry.

The trouble is that even the so-called experts might have fixed upon a description that fails to characterize a sufficiently explanatory kind. Imagine that all philosophers in fact associate ‘true belief formed in accordance with one’s evidence’ with the term ‘knowledge’; if descriptivism is correct, true beliefs formed in accordance with one’s evidence are knowledge. But none of this rules out the possibility that there is some other way of carving up epistemic states that would be more natural, more explanatory, more predictive, better suited to our theoretical purposes. Call the most natural, most explanatory/predictive category of epistemological states *shknowledge*. The trouble, then, is that descriptivism guarantees that intuition will reliably generate true beliefs.

---

60 Here and elsewhere, I assume that there is a unique set of optimally explanatory categories. This is primarily for simplicity’s sake; my position is also compatible with the view that there are multiple ways of carving up the world which are equally explanatory. I am only committed to the claim that SOME carvings are more suitable for theoretical purposes than others – i.e., not ALL carvings are equal.
about knowledge, but it does not give us any reason to believe that knowledge is

*shmowledge*.

The degree to which this observation expresses a real problem for a constitutive view of the reliability of intuition will depend upon what one takes the theoretical purposes of philosophy to be. There are at least two obvious options here. One the one hand, one could be, in Goldman and Pust’s (1998) terms, an ‘extra-mentalist’ – that is, one could take the targets of philosophical theorizing to be natural kinds, or at least external kinds ‘in the world’ whose natures do not depend upon our conceptions of them. It is this type of view on philosophical inquiry, I take it, that is expressed by e.g. Williamson’s insistence that the goal of metaphysics “is to discover what fundamental kinds of things there are… not to study the structure of our thought about them” (Williamson 2007, 19), or Sosa’s claim that “the question is not just whether ‘knowledge’ applies to the protagonist in a certain example. The question is whether the protagonist who satisfied the conditions specified in the example would *know*” (Sosa 2009, 104). More obviously, this is the view expressed by philosophers, like Devitt and Kornblith, who take philosophical investigation to be parallel to investigation of natural kinds.

The other obvious option is to claim that the aim of philosophy is to characterize our concepts, where ‘concept’ is understood in a psychological sense. This is the view that Goldman ascribes to. I will argue in the final section that this approach answers the concern of theoretical utility while maintaining constitutivity – at the cost of plausibly only capturing a small portion of actual philosophical practice. In the next section, however, I will claim that Jackson, while in some ways appearing to endorse a Goldmanian view, in fact makes several claims which are inconsistent with such a view
and which suggest instead an extra-mentalist approach. I’ll then argue that, if Jackson takes such an approach, the role his account allows for intuition will be quite minimal.

4. The extra-mentalist approach to philosophical theory

Jackson’s emphasis on folk theory and on the reactions of subjects to thought experiments appears, at first glance, to suggest that he endorses a mentalist approach to philosophical theorizing. However, upon closer examination, some of Jackson’s claims are in fact inconsistent with such an approach. To begin, Jackson permits rejecting a subject’s intuitions when faced with “signs of confused thinking”, “cases where their classification is… a derivative one”, or “readiness to back off under questioning” (Jackson 1998, p. 35). Perhaps this only suggests that Jackson aims to characterize conceptual competence, rather than conceptual performance. But further statements make this less plausible.

Jackson states that there is “nothing sacrosanct about folk theory” (Jackson 1998, 44). He notes that his intuitive conception of free will is one which is inconsistent with determinism; he then goes on to state that compatibilism provides a conception of free will that is “near enough to the folk’s to be regarded as a natural extension of it, and which does the theoretical job we folk give the concept of free action in adjudicating questions of moral responsibility and punishment” (Jackson 1998, 44-45). This sounds very much like an endorsement of an extra-mentalist take on the aims of philosophy.

Jackson also acknowledges, during the chapters in his book that are concerned with ethics, that variation in moral principles is rather prevalent, and claims that the
proper response to this variation is to see folk morality as a work in progress, whose principles and exceptions are currently under negotiation. What really determines what our analysis will be, Jackson claims, is *mature* folk morality – what the folk converge upon after sufficient rational reflection. Once we converge upon mature folk morality, presumably, the scope of our semantic knowledge will be sufficient to produce determinate application conditionals for every scenario. “The idea is that mature folk morality is the best we will do by way of making good sense of the raft of sometimes conflicting intuitions… that make up current folk morality” (Jackson 1998, 133).

If this strategy is permissible in the case of ethics, though, shouldn’t it be permissible elsewhere? Shouldn’t our analyses of mental states, for example, aim at the concepts of *mature* folk psychology? And, isn’t it likely that mature folk psychology would just be *-psychology*? What is the difference, here, between the folk reflecting on and potentially altering their own terms in light of theoretical considerations, and scientists or philosophers doing it for them?

Even by Jackson’s own lights, there may not be much difference. Consider the following statement Jackson gives about ‘theoretical terms’. Jackson writes that “many theoretical terms like ‘acid’, ‘kinetic energy’, ‘fish’, and ‘acacia’ are ones whose extension is in part determined by the nature of the best, true theory in which the term appears” (Jackson 1998, 34). Thus, the analysis we give for a term like ‘fish’ will be something like ‘x is a fish iff x has the important properties out of, or descended from, or explanatory of F1, F2, F3, according to the best true theory’, where F1, F2 and F3 are features we initially associate with fish exemplars (gills, scales, fins…). Such a description ensures that nothing will count as a fish unless it falls under a category
employed by a successful scientific theory of fishy entities – i.e., unless it falls under an explanatory kind. Given Jackson’s statement that it is in fact “an implicit part of serious classificatory practice that we seek to mark the divisions worth marking” (Jackson 1998, 35), one begins to suspect that a story like this might be required for philosophical terms as well. Henceforth, we’ll call this sort of strategy *explanatory descriptivism*.

Of course, there are some immediate complications with this strategy, as can be seen by considering terms like ‘jade’. Jackson’s description schema, when applied to ‘jade’, would seem to imply that ‘jade’ refers to the disjunction of jadeite and nephrite – for possession of the property ‘being jadeite or nephrite’ explains the presence of surface features that we initially associate with jade exemplars. Now, there is nothing wrong with such a conclusion from the standpoint of a theory of reference; many would even find it intuitive (for whatever that’s worth). But it is certainly not clear that such a category reflects a ‘division worth marking’, at least from a mineralogical perspective. Therefore, there appear to be cases where an explanatory descriptivist strategy fails to isolate an explanatory kind. For now, however, we can put this concern aside.

Jackson does not explicitly state which, if any, philosophically central terms count as ‘theoretical’. As noted above, however, certain of his statements do not mesh well with a purely psychological, concept-characterizing approach to the aims of philosophy. It will be worthwhile, then, to explore the consequences of embracing fully the explanatory descriptivist strategy for philosophical terms.
Thus, let’s assume that for terms like ‘knowledge’, speakers possess descriptive information along the lines of the schema given above for ‘fish’. Effectively, this involves specifying that something counts as knowledge iff it falls under what the best true epistemological theory takes to be the relevant explanatory kind. However, the trouble is that until we possess a true epistemological theory, this sort of a priori information about the nature of knowledge will not get us very far.

The point may become clearer when recast in terms of Jackson/Chalmers application conditionals. Jackson and Chalmers are fond of employing the Gettier thought experiment as an illustration of the importance of such conditional knowledge; we judge as we do on Gettier cases because we have knowledge of an a priori application conditional, G -> ~K, where G specifies the conditions laid out in Gettier’s thought experiment. The information in G, then, involves facts about the truth of the protagonist’s belief, its justificatory status, and its ‘accidental’ nature. However, for ‘theoretical’ terms, the information provided by the antecedent of the conditional must be much more complex. It must include information sufficient for determining whether the entities described fall under what the best true theory takes to be the relevant explanatory kind.

If explanatory descriptivism applies to ‘knowledge’, then the a priori application conditionals corresponding to ‘knowledge’ do not include, e.g., T -> ~K, where T is a description of the ‘Truetemp’ scenario that many take to be a counterexample to reliabilism, and K is the assertion that the mental state described in T is knowledge. Instead, they include the conditional (T&~E) -> ~K, where T and K are as stated above and E is the assertion that the mental state described in T falls under what the best true
epistemology (at that world) takes to be the relevant explanatory kind (at that world).

Unfortunately, however, the application conditionals associated with ‘knowledge’ also include (T&E) -> K! Essentially, if explanatory descriptivism is true for knowledge, then we know that if Truetime’s belief about the temperature falls under the relevant explanatory kind, then it is knowledge, and if it does not fall under the relevant explanatory kind, it is not. I hope it is clear that this analysis leads to almost no progress at all with regard to debates over reliability. If explanatory descriptivism is true, then the real bulk of the work of an epistemological theory must be in determining what the explanatory kinds are at this world – i.e., which antecedent actually holds. And there is no more reason to suppose that this can be determined solely by conceptual analysis than there is to suppose that chemistry might proceed solely by conceptual analysis. The ‘divisions worth marking’ are simply not given a priori.

There is some indication that Jackson might, in fact, agree with this assessment of the situation – at least in part. Jackson emphasizes that his view assigns conceptual analysis a ‘modest role’. Modest conceptual analysis tells us when, e.g., x counts as a belief in the folk sense of belief. It does not tell us anything about what the world is like – it does not, for instance, tell us that humans have beliefs. Nor can it tell us that a given scientific theory is false because it characterizes the brain in a way inconsistent with folk psychological theory. Modest conceptual analysis tells us what it takes to be an F, not whether there are any F’s or whether F’s play any role in a scientific understanding of the world.

Jackson does, however, often speak as though we must do conceptual analysis to avoid ‘changing the subject’:

“If I say that what I mean – never mind what others mean – by ‘belief’ is any information-carrying state
that causes subjects to utter sentences like ‘I believe that snow is white’, the existence of beliefs so conceived will be safe from the eliminativists’ arguments. But [I will not] have much of an audience. I have turned interesting philosophical debates into easy exercises in deductions from stipulative definitions together with accepted facts” (Jackson 1998, 31).

But it is my contention that, on an extra-mental approach to philosophical theorizing, this sort of stipulation would be an entirely appropriate way to proceed in philosophical debate – provided that the stipulator presented arguments in support of the claim that information-carrying states which cause belief reports comprise a more explanatory kind than any competitor.

Granted, there must be some connection between the folk concept and any proposed replacement concept – electrons comprise a strongly explanatory kind, but a proposal stipulating that ‘belief’ shall mean electron is inappropriate. This is, of course, because the explanatory role the electron plays is entirely unrelated to the role the folk notion of belief attempts to fill. But contra Jackson, there is no obvious reason to demand that, e.g., a philosophical account of knowledge need bear anything but the loosest of connections to the folk concept of knowledge. And clearly only the most minimal degree of ‘conceptual analysis’ is required to ensure that this sort of loose connection holds – a layman’s understanding of the folk term is surely sufficient. What is required is some degree of continuity in one’s theorizing; but no deep analysis of the initial concept is necessary to maintaining such continuity.

Elsewhere, in response to the idea that we might introduce a whole array of terms ‘belief₁’, ‘belief₂’, etc. and proceed to investigate for each whether it exists, Jackson writes:

“We have not actually addressed the issue on the table – the existence of belief without a subscript. We have ducked it… surely we do not want to end up in the position of refusing to discuss whether or not there is a cure for AIDS or was a big bang… while being prepared to discuss whether or not there is a cure₁ for AIDS₁ or was a big bang₅,” (Jackson 2009, 68).
On the contrary, however, I see no issue with being in such a position. If our theoretical aims are best served by introducing multiple notions to replace a single overbroad, disjunctive, or arbitrary category, then so be it.

I conclude that the Jacksonian approach does not vindicate anything like traditional philosophical methodology. The role of conceptual analysis in the characterization of explanatory kinds, on the explanatory descriptivism approach, is so minor as to be negligible; the role of conceptual analysis in current philosophical practice is anything but. The discussion in the paragraph above does, however, indicate that Jackson may yet have some sympathy for the mentalist approach to philosophical theorizing; indeed, the only sense I can make of the insistence on characterizing ‘belief without a subscript’ is that Jackson believes that an important goal of philosophy is to characterize the notion of (subscriptless) belief that is implicit in folk theory.

5. The psychological approach to philosophical theory

In this final section, I will examine the prospects for the mentalist, who takes the targets of philosophy to be ‘inside the head’. Alvin Goldman has explicitly endorsed a psychological version of the constitutivity approach to the defense of intuition. Specifically, Goldman takes it to be the case that “possessing a concept makes one disposed to have pro-intuitions toward correct applications and con-intuitions toward incorrect applications — correct, that is, relative to the contents of the concept as it exists in the subject’s head” (Goldman 2007, 15). In other words, intuition reveals psychological facts: my intuition that x is a case of knowledge is evidence –
constitutively-grounded evidence – that x falls under my knowledge concept. ‘Concept’ is here used in a narrow psychological sense, to indicate a particular mental representation whose content is fixed solely by what’s ‘in the head’ of the possessor. The reference of public language words is a separate issue – thus, Goldman need take no stand on descriptivism.

Goldman holds that although there is a general tendency for concepts to cause intuitions that accord with their contents – indeed, such a tendency seems to be part of what it is to have a concept in the first place – people can still sometimes make mistakes. We are occasionally subject to general performance related errors, such as poor memory, inattention, and so forth. In addition, we may have explicit theoretical beliefs about the nature of F’s, or about the nature of our own concept of F-hood, that override our pre-theoretic application intuitions. Thus, intuitions are a reliable, yet fallible, guide to the nature of personal psychological concepts.

Goldman claims that his account of the nature of intuition not only provides a naturalistically acceptable justification for intuition’s evidential status, but also squares well with some of the methodological quirks of philosophers. For instance, philosophers are typically wary of intuitions that have been tainted by theory; this makes a good deal of sense in the context of an attempt to characterize pre-theoretical personal psychological concepts. More generally, a psychological take on the aim of philosophy provides legitimacy to philosophers’ tendency to treat intuitions as evidential.

It may be argued that personal concepts cannot be the ultimate aim of philosophical inquiry; at the very least we must be concerned with public, shared concepts as well. Goldman agrees that personal concepts are not the sole aim of inquiry,
and discusses some ways in which we might use evidence about individual concepts to draw conclusions about shared psychological concepts. However, he comes to the conclusion that there is “no satisfactory way to promote a public or community-wide conception of concepts to the primary, or central, position in the project of conceptual analysis” (Goldman 2007, 18). Our primary focus must be individual concepts, with conclusions about shared concepts being a derivative aim.

On Goldman’s view, then, philosophers appeal to (e.g.) knowledge intuitions not with the intention of characterizing a suitably explanatory epistemological kind, but rather with the intention of characterizing a particular psychological feature of an individual. This information, in turn, might aid a more general project of psychological explanation, especially if at least some generalizations can be drawn about the nature of epistemological concepts across persons. Rather than drawing conclusions about the nature of some mind-independent phenomenon, then, philosophers might aim to draw conclusions about the ways in which certain groups (or even humanity as a whole) assign epistemological praise and blame, and how such tendencies affect behavior and cognition.

However, it is my contention that this sort of project, though certainly worthwhile in its own right, does not reflect the actual aims of many philosophers. Though Goldman is right in pointing out some ways in which his psychological take on intuition explains philosophical methodology, there are other features of philosophical practice that make little sense when viewed from the psychological perspective. For instance, philosophers are often willing to ‘bite the bullet’ when a particular feature of their account fails to square with intuition. Such a response would make little sense if the aim of philosophy
was to characterize one’s own concepts. Secondly, analyses of philosophical notions are often rejected when they are discovered to involve inconsistencies or other theoretical flaws like ontological extravagance. But there is no reason to expect that our personal psychological concepts respect such stringent demands – thus, philosophers interested in the psychological project should be happy to embrace inconsistent or otherwise inelegant analyses.

Further, if the dominant goal of philosophy were to characterize concepts in a psychological sense, then one would expect the current rise of experimental philosophy to be welcomed with open arms – after all, the methods of experimental philosophy seem to provide at least some insight into the concepts (in the psychological sense) of participants. Suffice it to say that experimental philosophy has not generally enjoyed such a welcome. Another indication that the psychological aim is not dominant: objections to experimental philosophy frequently invoke claims that subjects’ intuitions were not subject to sufficient reflection or theoretical critique, or that the subjects’ intuitions were irrelevant because they are not ‘experts’. Such objections don’t obviously square with a pure psychological account.

Finally, many philosophers explicitly state that their projects are not psychological. Timothy Williamson, for example, writes that “few philosophical questions are conceptual questions in any distinctive sense” (Williamson 2007, 3), and that “much contemporary metaphysics is not primarily concerned with thought or language at all. Its goal is to discover what fundamental kinds of things there are” (Williamson 2007, 19). As another example, Ernest Sosa writes that philosophical
disputes seem “to be disputes about something more objective than just a description or analysis of our individual or shared concepts” (Sosa 2007, 100).

In fairness, there are two quite plausible responses Goldman might give here. The first is to claim that the psychological approach does best characterize the aims of most philosophers, even if they do not recognize this themselves – indeed, even if they sometimes act in ways which are inconsistent with these goals. The other plausible response is to claim that the psychological approach should be the aim of philosophy, because it is only on this approach that the methods of philosophy emerge successful.

Perhaps the most plausible conclusion to draw, however, is that standard philosophical methodology involves a running together of mentalist and extra-mentalist projects. My own belief is that the extra-mentalist project is generally more interesting and important; however, I find no fault in those who prefer to pursue the psychological project, so long as they do not take their psychological conclusions to entail parallel conclusions about the nature of extra-mental phenomena.

Indeed, the two projects might even have a synergistic effect. I emphasized in the last section that, on the extra-mentalist view, substantive analysis of folk concepts is not necessary to the project of delineating explanatory kinds - we do not need deep analysis of folk concepts in order to avoid ‘changing the subject’. However, the psychological project might nonetheless inform the explanatory project in various ways - particularly if the psychological project focuses on an understanding of the particular cognitive mechanisms behind folk concepts. Studies of the workings of folk concepts might suggest ways in which intuitive methods for carving up the world go wrong; rather than ensuring that our post-theoretic concepts match up well with our pre-theoretic ones, then,
psychologically-oriented conceptual analysis might provide the explanatory project with error theories.
Chapter Six

Epilogue

One of the most distinctive features of philosophical inquiry – from Plato through the present day – has been its reliance on the spontaneous, apparently non-inferential judgments commonly referred to as ‘intuitions’. Over the past fifteen years, this aspect of philosophical methodology has come under heavy scrutiny, and doubts have been raised about the widespread assumption that intuitions can constitute evidence for or against philosophical hypotheses. My aim in this dissertation has been to examine several common themes of this recent methodological debate which I take to be problematic, in the hopes of clarifying some of the conceptual territory surrounding these methodological debates.

Many of the concerns I have attempted to raise involve, to my mind, oversimplifications - particularly concerning the relationship between the reliability of our intuitions and their evidential status. I argued, first, that philosophers on both sides of the intuition debate have, tacitly or explicitly, embraced two claims about philosophical methodology which sit poorly together - that intuition forms a single, reasonably unified mental capacity, and that the epistemological status of this capacity will depend at least in large part upon its reliability. An examination of the psychological and neuroscientific literatures on intuitive judgment suggests that intuition is highly heterogeneous; and that, therefore, an epistemological assessment of intuition based on reliability may be overbroad.
I also argued that, while reliability is clearly an important consideration in the epistemological evaluation of the states falling under the term ‘intuition’, there is a frequent assumption that anyone who questions the suitability of intuitions as evidence must deny that intuitions are reliable enough to constitute knowledge. Thus, the anti-intuitionist is frequently characterized as claiming that intuitions produce massive amounts of error in judgment – which, if Williamson is correct about the inseparability of intuition from concept application generally, may threaten radical skepticism. I argued, however, that anti-intuitionists are not committed to a claim that intuition is subject to massive error. Indeed, it is perfectly consistent with the findings of various experimental studies on variation in intuition to claim that intuition is, on the whole, reliable. There is a fairly reasonable chance that our ‘intuitive’ judgments are on the whole reliable for the purposes of day-to-day classification, but that the degree of error they do exhibit makes them unsuitable for the purposes of the construction of philosophical theories. Should this be the case, an anti-intuitionist can successfully critique the use of intuitions as evidence without thereby risking skepticism.

A second theme of the dissertation has been the relationship between the categories delineated by our intuitive classifications and the categories which serve as the targets of philosophical investigation. If the cases we tend to classify as instances of knowledge (e.g.) can be somehow guaranteed to more or less line up with the cases that philosophers interested in knowledge aim to study, then this bodes well for the use of intuitions in philosophy. One way in which this might occur, of course, is if our intuitions must necessarily be for the most part true – i.e., if our classificatory judgments about knowledge must accurately reflect the extension of the concept knowledge. The
aim of constitutive approaches to the defense of intuition is to argue that this is necessarily so, given facts about the nature of the meanings of philosophical terms.

I argued that none of the most prominent versions of the constitutivity argument justify the current widespread use of intuition as primary evidence for philosophical theories. Against Bealer, I argued that self-defeat arguments alone cannot support his claim that philosophy possesses autonomy and authority; I further argued that the semantic features of philosophical terms do not obviously differ from those of natural kind terms in a way that would underwrite the reliability of intuition.

Against Jackson, I argued that a broadly descriptivist approach to the necessary reliability of intuition fails to guarantee that the referents of our folk concepts play an important role in philosophical theorizing. If we take an extra-mentalistic view on the aims of philosophy, the most plausible implementation of Jackson’s view involves what I have called explanatory descriptivism; however, that implementation leaves only a very minimal role for intuition. If we instead side with Goldman and take a mentalistic view on the aims of philosophy, intuition clearly has a greater role to play; further, the theoretical utility of folk classification is made clear – such classifications inform the project of characterizing concepts in the psychological sense. However, it is my contention that the mentalistic view of philosophy reflects only a minority of actual philosophical activity.

Ultimately, though I have provided no positive account of the use of intuition in this dissertation, I am attracted to a moderate position according to which certain instances of ‘reliance on intuition’ are acceptable, while others must be accorded far less evidential weight, and still others may need to be abandoned entirely. This will not entail the downfall of the discipline of philosophy, nor will it necessarily involve abandonment
of traditional *a priori* philosophical tools such as the thought experiment or the method of
counterexamples. But there is a pressing need for thorough examination and testing of
those methods, which will almost certainly result in some modification and restriction of
their use.
REFERENCES


University Press.


