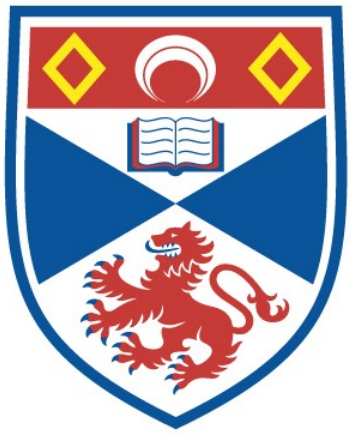
Intuition, metacognition, and philosophical inquiry

Miguel Egler

A thesis submitted for the degree of PhD
at the
University of St Andrews



2019

Abstract

This thesis examines whether appeals to intuition provide warrant for philosophical positions and claims. Recent empirical studies suggest that people's intuitions are sensitive to a host of epistemically irrelevant factors. Some philosophers (Critics) take these findings to motivate scepticism about the traditional methodology of philosophy, which is often taken to rely on intuitions. In reply, others (Defenders) have argued that these empirical studies fail to support any such methodological worries. One type of argument given by Defenders is that empirical studies about intuitions are simply irrelevant for assessing the methodology of philosophy. A second type of argument is that empirical studies need to be refined before they can be fruitfully used in debates about the methodology of philosophy: for instance, some Defenders have suggested that empirical studies must account for the distinctive phenomenology of intuitions; whereas others have argued that these studies should focus exclusively on the intuitions of philosophers—which, they claim, are not prone to the deleterious effects found to afflict the intuitions of non-philosophers.

In this thesis, I argue that the most prominent arguments both for and against scepticism about the methodology of philosophy are problematic. I defend instead a moderate position according to which intuitions *can* warrant philosophical positions and claims, but that philosophers should engage with findings in cognitive psychology in order to make better informed assessments of when to trust their intuitions, and when to refrain from relying on them in inquiry. In chapter 1, I argue that empirical findings about intuitions are indeed relevant for assessments of the traditional methodology of philosophy. In chapter 2, I appeal to recent work in cognitive psychology on the feeling of rightness to argue that a phenomenological conception of intuition does not insulate philosophical methodology from the relevant empirical concerns. In chapter 3, I argue that philosophical training can in *some* cases give rise to a kind of expertise that shields philosophers from the deleterious effects of biases on their intuitions; but, I contend that further empirical studies are required to establish what those cases are. And against Critics, I argue in chapter 4 that even if intuitions show marked variation with respect to epistemically irrelevant factors, this does not entail that we should be sceptical about their use in philosophical inquiry. Lastly, in chapter 5 I explore the proposal that intuitions provide warrant for philosophical positions and claims because they enable understanding of why their contents are true.

Acknowledgements

First and foremost, I would like to thank Jessica Brown for invaluable guidance and encouragement throughout all these years. I have been very fortunate to work with Jessica during this time, and to have benefited from her remarkable insight and sensitivity with the topics I was investigating. Jessica's incisive comments and criticisms have helped me to appreciate the high degree of intellectual rigour, attention to detail, and care required in philosophical investigation. For all this, and for her kindness and patience, I am very grateful to her. I would also like to thank Simon Prosser, who has been my secondary supervisor throughout all these years. I have greatly enjoyed meetings with Simon, which were always very encouraging and made me think harder and deeper about my research. I greatly appreciate his generosity with his time, and the vital feedback on my work.

Many thanks to all of those with whom I have discussed the material of this thesis throughout all these years (apologies for any omissions): Derek Ball, Lisa Bastian, Joe Bowen, Yuri Cath, Elijah Chudnoff, Michael Hannon, Jonathan Ichikawa, Bruno Jacinto, Ethan Landes, Lorenzo Lazzarini, Joachim Horvath, Ole Koksvik, Stefano Lo Re, Poppy Mankowitz, Matt McGrath, Jennifer Nado, Quentin Pharr, Joëlle Proust, Saranga Sudarshan, Janis Schaab, Jonathan Schaffer, Justin Snedegar, Erika Torres, Clotilde Torregrossa, Brian Weatherson, and Jonathan Weinberg. Many thanks to anonymous reviewers from *Mind & Language* for comments on chapter 2, and the anonymous reviewers from *Synthese* for comments on the material in chapter 3. Special thanks to Eugen Fischer, for continued support and encouragement, and for pointing me in the right direction. And very special thanks to Marvin Backes, Claire Field, and especially Lewis Ross, for the many discussions, comments, and criticisms that have helped me to hone my skills as a philosopher.

Many thanks to the whole Arché community. Arché really is a special and exciting place to do philosophy. I have learned so much from the countless discussions with so many brilliant and inspiring philosophers. It has really been a privilege to be around for the many conferences, seminars, reading groups, and other events that the centre has hosted throughout all these years. I will dearly miss being part of such an amazing group. Thanks also to Lynn Hynd for all her administrative support during this time.

Lastly, very very special thanks to Sophia and Moysés (couldn't have done it without you guys). To Marina, for everything. And to my parents, for even more.

Chapter 2 is forthcoming in *Mind & Language* as *Testing for the Phenomenal: Intuition, metacognition, and philosophical methodology*. Chapter 3 is forthcoming in *Synthese* as *Philosophical Expertise Under the Microscope*.

Candidate's declaration

I, Miguel Egler, do hereby certify that this thesis, submitted for the degree of PhD, which is approximately 62,600 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for any degree.

I was admitted as a research student at the University of St Andrews in September 2015.

I received funding from an organisation or institution and have acknowledged the funder(s) in the full text of my thesis.

Date Signature of candidate

Supervisor's declaration

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date Signature of supervisor

Permission for publication

In submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand, unless exempt by an award of an embargo as requested below, that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that this thesis will be electronically accessible for personal or research use and that the library has the right to migrate this thesis into new electronic forms as required to ensure continued access to the thesis.

I, Miguel Egler, confirm that my thesis does not contain any third-party material that requires copyright clearance.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

Printed copy

Embargo on all of print copy for a period of 3 years on the following ground(s):

- Publication would preclude future publication

Supporting statement for printed embargo request

The material of this thesis provides a novel contribution to the field, which I am currently preparing for submission to journals. I would like to request an embargo on this material so that this material is not made available to the public before the reviewing process of this material is finished.

Electronic copy

Embargo on all of electronic copy for a period of 3 years on the following ground(s):

- Publication would preclude future publication

Supporting statement for electronic embargo request

The material of this thesis provides a novel contribution to the field, which I am currently preparing for submission to journals. I would like to request an embargo on this material so that this material is not made available to the public before the reviewing process has finished.

Title and Abstract

- I agree to the title and abstract being published.

Date Signature of candidate

Date Signature of supervisor

Underpinning Research Data or Digital Outputs

Candidate's declaration

I, Miguel Egler, hereby certify that no requirements to deposit original research data or digital outputs apply to this thesis and that, where appropriate, secondary data used have been referenced in the full text of my thesis.

Date Signature of candidate

Contents

Introduction

This thesis examines whether appeals to intuition provide warrant for philosophical positions and claims. My conclusion is a resounding 'yes-and-no'. My argument for this position does not amount to a mere compromise between starkly opposing views; rather, in the chapters that follow I demonstrate that empirical and theoretical considerations provide strong support for this moderate position.

In this introduction, I provide an overview of recent metaphilosophical debates concerning the nature and epistemology of intuitions. I first discuss how philosophers appeal to intuitions in philosophical inquiry. I then detail the major developments in the innovative field of experimental philosophy, and explain how findings from studies in this movement putatively raise doubts about the methodology of philosophy. I then survey the most prominent objections to these experimentally-motivated methodological challenges and responses to these objections. This survey will outline the aims of this thesis.

0.1 Intuitions

0.1.1 Intuitions in philosophy

Intuitions are widely regarded as playing a central role in the traditional methodology of philosophy. The intuitions in question concern people's verdicts on whether a philosophical notion applies to a given (hypothetical or actual) scenario. Philosophers often make use of thought-experiments to elicit intuitions, which they then invoke in philosophical argumentation—generally in accordance with the principle that theories, positions, and claims that are consistent with those intuitions should be preferred over those that are not (see, e.g., Bealer, 1998; Pust, 2000). As an example, consider the following famous dialogue from Plato's Republic:

> A fine sentiment, Cephalus, but, speaking of this very thing itself, namely, justice, are we to say unconditionally that it is speaking the truth and paying whatever debts one has incurred? Or is doing these things sometimes just, sometimes unjust? I mean this sort of thing, for example: Everyone would surely agree that if a sane man lends weapons to a friend and then asks for them back when he is out of his mind, the friend shouldn't return them, and wouldn't be acting justly if he did. Nor should anyone be willing to tell the whole truth to someone who is out of his mind.

> That's true.
>
> Then the definition of justice isn't speaking the truth and repaying what one has borrowed.
>
> (I, 331 c-d)

In this passage, Socrates first describes an imaginary scenario and reports what he takes to be the widely-shared and correct intuitive verdict to it (namely, that one should not return the weapon). Then, in line with the principle that an adequate account of justice must accommodate for this intuitive verdict, he contends that the account of justice his interlocutor (Cephalus) advances must be false.

Beyond such a historical illustration, examples of appeals to intuition abound in contemporary debates. For instance, consider recent disputes surrounding the so-called 'Trolley problem'—i.e., a set of hypothetical scenarios where people are asked if it would be morally permissible to alter the course of a run-away trolley to save many people at the cost of sacrificing the life of one person (Thomson, 1985). Typically, theories about the moral permissibility of actions are required to accommodate for the fact that many people judge it morally permissible to change the course of the trolley—under the premise that this intuitive verdict is somehow representative of moral truth (for a discussion: Kamm, 2015). Furthermore, a quick survey of the philosophical literature provides a host of further instances of appeals to intuitions in contemporary debates within Epistemology, Philosophy of Mind, Philosophy of Action, and Philosophy of Language. Commenting on the prevalence of this intuition-based approach to philosophical inquiry, Janet Levin (2005, p. 194) claims that the "procedure of rejecting or modifying theses in the face of intuitively convincing counterexamples has been characteristic, perhaps definitive, of philosophical argumentation throughout its history". Alvin Goldman (2007, p. 1) suggests further that what "distinguishes philosophical methodology from the methodology of the sciences is its extensive and avowed reliance on intuition."

Now, despite widespread agreement that intuitions play an important role in traditional methods of philosophy,[1] there is little to no agreement on what precisely counts as an intuition. Indeed, a survey of the recent literature reveals a motley of mostly incommensurable views on this issue.[2] For the sake of simplicity, we can organise these accounts on a scale ranging from more minimalist views—which account for intuitions in terms of other more familiar mental states—to more robust views—which conceive of intuitions as a special mental state with a set of unique properties (e.g., a particular phenomenology or etiology). For the remainder of this section, I provide a brief description of some of the main positions on recent debates about the nature of intuitions.

Before proceeding, there are two observations that are here noteworthy. First, the discussion that follows is in no way comprehensive. My aim is not to outline and review *all* extant positions in this debate; rather, I focus here on those positions that have proven most influential and which are germane to the issues that I address in this thesis. Second, I will not seek to adjudicate between any of these positions. Indeed,

[1] Some famous dissenters from this proposal include Herman Cappelen (2012) and Max Deutsch (2015). I discuss their views below (sec. 0.4.1).

[2] For a survey, see Cappelen (2012, pp. 1-23)

throughout this thesis I will remain neutral on the question of what is the correct view of the nature of intuitions. With these considerations in mind, I now turn to the discussion.

0.1.2 The nature of intuitions

Minimalist views take intuitions to be entirely accountable for in terms of other more familiar doxastic propositional attitudes—such as beliefs, judgements, and/or inclinations to believe or judge (Lewis, 1983; Inwagen, 1997b; Williamson, 2007; Earlenbaugh and Molyneux, 2009). Williamson has offered perhaps the most detailed and influential defence of such a minimalist conception of intuition. His argument for this position builds on the observation that the term 'intuition' is used quite liberally in philosophy. Indeed, as he points out, philosophers have denominated almost all kinds of judgement—even straightforward perceptual ones—as 'intuitions'. Williamson proposes that, given this indiscriminate (or, as he puts it, 'promiscuous') use of the term, we should adopt a fairly broad conception of intuitions so as to not rule out any such uses of 'intuition' by fiat. Thus, he suggests that intuitions are not any *special* kind of mental state, but just ordinary beliefs, judgements, or inclinations to believe or judge.

Robust views of intuitions, on the other hand, suggest that attention to the particular features of intuitions shows them to be irreducible to other more familiar mental states. We can divide the robust views in two broad camps: phenomenalist and etiological views. I explain each in turn.

Phenomenalist views regard intuitions to be a particular kind of mental state with a characteristic phenomenology. Broadly speaking, these views all defend the claim that intuitions make their contents *seem* true to a subject. This proposal requires further elucidation. For this end, Bealer's claims are here instructive:

> For you to have an intuition that A is just for it to seem to you that A. Here 'seems' is understood [...] in its use as a term for a genuine kind of conscious episode. For example, when you first consider one of de Morgan's laws, often it neither seems to be true nor seems to be false; after a moment's reflection, however, something new happens: suddenly it seems true.
>
> (Bealer, 1998, p. 271)

Phenomenalists have taken great pains to clarify just what it means to say that an intuition makes its content *seem* true in the way that Bealer describes (see, e.g., Pust, 2000; Bengson, 2015b; Chudnoff, 2013; Koksvik, 2011).[3] For this aim, they often appeal to a historically influential "perceptual analogy": that is, the idea that intuitions and perceptual experiences are mental states that share intrinsic commonalities. Phenomenalists interpret this analogy in more narrow terms as the claim that *what it feels like* to have an

[3] There are many intricate and complex differences amongst all these accounts. For instance, some phenomenalists impose further restrictions on what counts as an intuition—e.g., they must have a specific kind of content, or would have that content under specific circumstances (e.g., Bealer, 1998; Pust, 2000, p. 38). For current purposes, we can ignore these intricate qualifications of the phenomenalist views.

intuition is similar to *what if feels like* to have a perceptual experience. For example, they suggest that the phenomenal contours of an intuition which makes it *seem* that triangles have three sides bears important commonalities with the phenomenology of a perceptual experience that makes it *seem* that one sees a red light shining in the distance. Building on such illustrations, phenomenalists argue that intuitions are best characterised as involving an *intellectual* seeming that is similar (but not equivalent) to a perceptual seeming.

Phenomenalists argue further that on this perceptual view, it becomes clear that intuitions are not reducible to beliefs, judgements, or inclinations to believe or judge. Their main contention is that one can have an intellectual seeming with respect to a proposition, yet not believe, judge, or be inclined to believe or judge that it is true. For instance, many have reported having an intellectual seeming (an intuition) that makes the naïve comprehension axiom seem true to them, even though they know that this axiom is false (and thus, do not believe nor judge it to be true). Taking this example as illustrative, phenomenalists contend that minimalist views fail to adequately account for the nature of intuitions.

A second prominent set of robust views are etiological conceptions of intuitions. Proponents of etiological views claim that there is a particular mechanism generating intuitions that distinguishes them from other mental states. One of the most prominent etiological views is that intuitions are mental states that stem solely from our *conceptual competence*. Kirk Ludwig provides a clear articulation of this proposal in the following passage:

> For terminological clarity, I will use "intuition" to mean an occurrent judgement formed *solely* on the basis of competence in the concepts involved in response to a question about a scenario, or simply an occurrent judgement formed solely on the basis of competence in the concepts involved in it (in response, we might say, to the null scenario). Intuitions in this sense are what we seek to elicit in response to questions about scenarios in thought experiments.
>
> (Ludwig, 2007, pp. 135-136)

The relevant notion of 'conceptual competence' at stake here concerns a capacity to deploy concepts in a way that *accurately* reflects their application conditions. In effect, this means that on such etiological views, intuitions are mental states that express conceptual *truths*. Proponents of conceptual competence views argue that it is thereby clear that not every judgement can be regarded as an intuition. After all, people often hold beliefs or make judgements that reflect mistaken use of concepts (when, e.g., they make hasty evaluations, or are misinformed about how to use that concept). On this note, intuitions are taken to not be reducible to an ordinary doxastic propositional attitude (ibid.; for similar positions: Kauppinen, 2007; Grundmann, 2010).

This completes the brief survey of different views of the nature of intuitions. Again, it is crucial to emphasise that this is not a comprehensive overview of all major positions in the debate.[4] And neither

[4] For instance, I have not here discussed another increasingly prominent etiological view according to which intuitions are

does this purport to show that any such position is in better standing. My aim in outlining points of contention amongst these accounts is that they help to highlight some of the topics that I address in this thesis.

0.2 Experimental philosophy

The term 'experimental philosophy' refers to a broad movement that seeks to employ methods and findings from the empirical sciences to shed light on philosophical issues and questions.[5] Thus defined, experimental philosophy is not revolutionary: Rather, it is best characterised as a natural extension of projects in early philosophical naturalism (Rose and Danks, 2013). However, over the last 20 years 'experimental philosophy' has been used more narrowly to refer to empirical investigations of philosophically-relevant *intuitions*. It is this sense of 'experimental philosophy' that I will focus on in this thesis.

The methods and approaches of experimental philosophy fall into roughly three broad camps. The first involves the use of questionnaire-based surveys to investigate whether people's intuitions correlate with certain target variables. A typical example is the use of questionnaires to test whether people from a given demographic group (e.g., from a given culture or age group) tend to have a specific intuition in response to a philosophical thought-experiment (e.g., that the subject in Gettier cases does not know). The second concerns the use of experiments to test whether intuitions are sensitive to the manipulation of certain defined parameters. One example of such approach involves dividing participants into two groups where participants read the same set of cases, but each group reads them in a different order. Results from this experiment can then track whether participants in these groups have distinct intuitions about these cases depending on which they read first. The third concerns the use of empirical findings and methods to provide psychological explanations of intuitions. Generally, this approach involves the use of research in cognitive psychology with the aim of pinpointing the cognitive processes underwriting philosophically-relevant intuitions. These examinations thus afford an improved understanding of the sources of philosophically-relevant intuitions.

Philosophers have made use of these methods and approaches for distinct purposes. We can better define these distinct applications by listing the current three major projects experimental philosophy. These are:

The Conceptual Analysis project: One of the main projects in experimental philosophy seeks to use empirical methods to aid efforts in conceptual analysis. Broadly stated, methods of conceptual analysis make recourse to people's intuitions about the deployment of a given concept in order to develop an analysis of its application conditions. Some experimental philosophers claim that techniques and methods from experimental philosophy are useful for this end. Indeed, some have argued further that

defined by the cognitive processes that give rise to them (see, e.g., Nagel, 2010; Nagel, 2012b; Fischer, 2014; Alexander, Gonnerman, and Waterman, 2015; Gerken and Beebe, 2016; Fischer and Engelhardt, 2016; Fischer and Engelhardt, forthcoming). I have also not here discussed some accounts that do not neatly fit into the distinctions I have traced here—such as Sosa's (2007b) proposal, which is both etiological and phenomenalist.

[5] For recent collections, see Fischer and Collins (2015) and Sytsma and Buckwalter (2016).

experimental approaches are in fact better suited for the aims of conceptual analysis, when compared to non-experimental (armchair) theorising (see, e.g., Stich and Tobia, 2016). The central motivation for this claim is that empirical methods allow for a detailed and large-scale examination of the intuitions of a large group of people, rather than of just those of a restricted group of philosophers. To testify for the effectiveness of this approach, experimental philosophers usually highlight empirical studies that have uncovered surprising patterns in how people apply certain philosophically-relevant concepts. For instance, one prominent study in experimental philosophy has found that people's attributions of intention are strongly linked to moral assessments—a pattern of evaluation denominated the "side-effect effect" (Knobe, 2003). Evidence that people's evaluations were prone to this "side-effect effect" was a surprising discovery for many philosophers. Moreover, the fact that this effect has escaped so many philosophers—despite decades of research in the philosophy of action and moral philosophy—has been taken to show the advantages of experimental techniques for projects in conceptual analysis.

The Cognitive Science project: Many recent studies in experimental philosophy are aptly characterised as integrating research efforts in Cognitive Science (Knobe, 2016). These studies make recourse to experimental techniques and questionnaire-style surveys to identify the specific cognitive processes that generate important philosophical intuitions. The primary aim of such investigations is simply to gain a better understanding of the workings of these intuitions and of the mechanisms responsible for some of the notable patterns found to influence them. One example of such projects are recent inquiries into the cognitive processes underwriting the side-effect effect—that is, the tendency to take moral assessments in consideration when attributing intention to others (discussed above). Much recent work on the side-effect effect has sought to identify the particular cognitive processes responsible for generating intuitions that display this pattern, and to investigate whether they underwrite other philosophical intuitions as well (Alicke, Gordon, and Rose, 2013; Knobe and Szabó, 2013; Newman, Bloom, and Knobe, 2014; Turri and Blouw, 2015).

The Warrant Project: Another major project in experimental philosophy aims to make use of empirical methods to assess whether intuitions provide warrant for philosophical positions and claims. We can divide studies in the Warrant Project into two camps:

The Variation Project: Studies in the Variation Project aim to elicit intuitions by way of questionnaire-style surveys or experiments, and seek to either expose conflicts between (a) the intuitions of philosophers and non-philosophers, or (b) determine whether people's intuitions are sensitive to certain defined parameters—such as their demography, the order in which they read each case, or the framing of the relevant questions. These studies aim to derive normative assessments of people's intuitions directly from these conflicts or sensitivities (for a review: Machery, 2017, ch. 2).

Cognitive Epistemology: Studies in Cognitive Epistemology aim to provide psychological explanations of intuitions that can help to assess their reliability. These studies first trace intuitions back to a specific cognitive process. Then, they make use of empirical findings about the functioning of that cognitive process to identify whether they are likely to produce reliable outputs, and under what circumstances they are likely to lead judgements astray. Relying on such analyses, studies in Cognitive Epistemology seek to

provide assessments of whether that particular intuition should pull any weight in philosophical theorising (Nagel, 2010, e.g., Nagel, 2012b; Fischer, 2014; Alexander, Gonnerman, and Waterman, 2015; Gerken and Beebe, 2016; Fischer and Engelhardt, 2016; Fischer and Engelhardt, forthcoming)).

The Warrant Project has proven to be, by far, the most influential project in experimental philosophy. This is due to the surprising number of studies suggesting that people's intuitions are sensitive to a host of epistemically irrelevant factors. Some philosophers have argued that these studies in the Warrant Project raise serious concerns about traditional methods of philosophical inquiry, which are taken to often rely on intuitions. This experimentally-motivated methodological challenge has come to be known as the 'experimentalist challenge'. In the next section, I briefly discuss the body of findings which seem to suggest that intuitions are prone to such deleterious effects, and further explain why they putatively motivate the experimentalist challenge.

0.3 The experimentalist challenge

The findings that putatively motivate the experimentalist challenge can be helpfully divided into two categories. The first comprises evidence that people's intuitions can vary along demographic lines. I will denominate these as findings of 'cognitive diversity'. Of all the findings in experimental philosophy, it is the evidence of cognitive diversity that has received the most attention. This is especially the case for findings that seem to show cross-cultural differences in people's intuitions about philosophical cases. However, other studies seem to suggest that people's intuitions also vary with respect to their personality, gender, and age. And, although a great number of such studies have recently failed to replicate (see, e.g., Nagel, San Juan, and Mar, 2013; Adleberg, Thompson, and Nahmias, 2015; Kim and Yuan, 2015; Seyedsayamdost, 2015a; Seyedsayamdost, 2015b; Freitas et al., 2018), there is growing evidence that at least some intuitions may actually be demography-specific (see, e.g., Machery et al., 2004; Feltz and Cokely, 2012).

A second body of findings motivating the experimentalist challenge indicate that people's intuitions can vary depending on how philosophical questions are presented to them. I dub these as findings of 'intra-personal variation'. Most studies reporting these findings have tested for two well-known patterns of variation from the literature on cognitive biases. The first are called "framing effects": a pattern of evaluation in which people's answers vary depending on the wording of the task at hand. For example, people have been found to be more likely to judge that a medication was good when were told it was 50% effective, when compared to those who were told that it fails 50% of the time (Levin, Schnittjer, and Thee, 1988). The second are "order effects": a tendency to provide different evaluations to a set of questions depending on the order in they are presented. For instance, in one experiment participants were first asked how satisfied they were with their current situation as a whole, and subsequently how satisfied they were with their marriage. In this ordering of the questions, people's ratings of satisfaction with their lives and their marriage were only weakly correlated; but, when the question order was inverted, this correlation was significantly increased. Researchers take these findings to show that priming people to think about their marriage led them to measure their degree with satisfaction with their lives as a function

of their satisfaction with their marriage (Schwarz, Strack, and Mai, 1991). Some studies in experimental philosophy make use of similar techniques to investigate if people's intuitions are prone to such effects. Surprisingly, many of these studies report robust evidence of intra-personal variation of order effects and framing effects in intuitions about prominent philosophical thought-experiments (for a review: Machery, 2017, ch. 2).

The above findings suggest that there is something amiss with traditional methods of philosophy, which is often taken to rely on intuitions.[6] But, these findings indicate that people's intuitions are sensitive to, inter alia, their demography, as well as the order and the wording of philosophical cases. Given the sensible assumption that none of these factors should play *any* role in philosophical debates, it then seems reasonable to conclude that the methodology of philosophy needs to undergo revisions so as to rule out these distorting influences in philosophical theorising. We can express this argument more schematically as follows:

The experimentalist challenge

P1 Intuitions play a central role in traditional methods of philosophy.

P2 Experimental findings reveal that people's intuitions are sensitive to epistemically irrelevant factors.

P3 If an intuition is sensitive to an epistemically irrelevant factor, then it is unsuitable for use in philosophical inquiry.

C Traditional methods of philosophy stand in need of revisions.

Note that the conclusion of the experimentalist challenge does not specify what kinds of revisions are warranted by the relevant empirical findings. Proponents of the experimentalist challenge are divided on this issue. We can identify two major positions in this dispute. I explain each in turn.

On more radical versions of the experimentalist challenge, the relevant empirical findings provide decisive reasons to conclude that intuitions provide no warrant for philosophical positions and claims. These findings are thus taken to motivate scepticism about traditional methods of philosophical inquiry, which are assumed to often rely on appeals to intuitions. Generally, defences of this position accompany a strong sceptical stance with respect to philosophical views that rely on appeals to intuition. Stephen Stich provides an exceptionally clear articulation of this idea. He proposes that *if* there really is cognitive diversity with respect to philosophical intuitions, "then a great deal of what goes on in contemporary philosophy, and a great deal of what has gone on in the past, belongs in the rubbish bin" (Stich, 2009, p. 232). Machery has recently expressed similar thoughts, claiming that evidence of cognitive diversity

[6] There are other empirical findings beyond these listed here that *could* also be taken to motivate the experimentalist challenge. For instance, some findings in Cognitive Epistemology (described in the previous section) suggest that intuitions can be generated by cognitive processes that are unreliable. These findings could be taken to motivate experimentally-motivated concerns about the methodology of philosophy. But, proponents of the experimentalist challenge have not made recourse to these findings when articulating these methodological concerns. As such, I will not discuss these findings in this context.

provides decisive reasons to think that philosophers should abandon projects that rely on intuitive verdicts about philosophical thought-experiments (Machery, 2017, p. 145). Swain and colleagues express similar concerns: they suggest that the evidence of intra-personal variation in philosophical intuitions "make live and salient the possibility that [proponents of traditional methods of philosophy] will find that their practice may in fact be built on an unacceptably shifting foundation" (Swain, Alexander, and Weinberg, 2008, pp. 153-154).

Proponents of moderate versions of the experimentalist challenge object to these strong conclusions. To clarify their proposal, it is here useful to consider an analogy with the role of perception in scientific investigations. Notably, there is considerable evidence that trusting our perceptual faculties can on occasion lead us to make erroneous judgements. Examples include cases of optical illusions, as well as instances where dim lighting can be deceiving. Now, evidence that perception can lead judgements astray in these ways plausibly does not show that perception fails to provide warrant for scientific claims, nor does it motivate scepticism about its use in scientific inquiry. Instead, the reasonable conclusion to draw from this body of evidence is that scientists ought to take measures to rule out that their perceptual faculties are prone to these problematic influences when appealing to them in scientific inquiry. Proponents of moderate versions of the experimentalist challenge make recourse to similar considerations when discussing the methodological implications of evidence that intuitions are sensitive to epistemically irrelevant factors. On their view, the relevant empirical findings do not preclude that intuitions can provide warrant for philosophical positions and claims; but they take them to show that philosophers need to develop methods for identifying problematic intuitions so as to avoid appealing to them in philosophical inquiry (Alexander and Weinberg, 2014; Weinberg, 2015; Nado, 2016b; Weinberg, 2016). Moreover, they propose that experimental philosophy can prove useful for this end. Their central claim is that findings about the workings of intuitions effectively pinpoint sources of error in philosophical intuitions; and, as such, they can help philosophers make better informed assessments of when to trust their intuitions, and when to refrain from relying on them in inquiry. In this sense, they contend that experimental philosophy provides valuable resources with which to *improve* traditional methods of philosophy.

At this point, it is important to pause and make an important clarification about the discussion above. In particular, it is noteworthy that most proponents of the experimentalist challenge assume that intuitions *in general* enjoy a default positive epistemic status (Weinberg, 2007; Stich, 2009, p. 242; Alexander and Weinberg, 2014; Machery, 2017, p. 183).[7] That is, they regard most of our ordinary intuitions about, say, semantic reference, causation, knowledge, or justification to be warranted. However, they contend that empirical findings about the workings of intuitions raise worries about their use for the purposes of philosophical inquiry. In this sense, the crux of the experimentalist challenge is the claim that the relevant empirical findings motivate worries about what philosophers *do* with intuitions (and not so much a worry about the epistemic status of intuitions *taken as a whole*). On radical formulations of the experimentalist challenge, the relevant findings provide decisive reasons to think that intuitions lose their default positive epistemic status *in the context of philosophical inquiry*—and that philosophers should therefore refrain

[7] Arguably, Cummins (1998) can be taken to deny this claim.

from appealing to them in that context. By contrast, moderate formulations of this challenge take these findings to show only that philosophers should make recourse to experimental methods and findings in order to guide their use of intuitions in inquiry.

Throughout this thesis, I argue for a moderate version of the experimentalist challenge. My argument for this position is two-fold. First, I aim to show that attempts by the defenders of traditional methods of philosophy to undermine the experimentalist challenge fail (call proponents of these views 'Defenders'). Second, I argue that attempts to establish a radical versions of the experimentalist challenge also face significant difficulties (I dub proponents of these views 'Critics'). My arguments against Defenders and Critics will underscore the lack of sufficient reasons to deny that intuitions can provide warrant for philosophical positions and claims, whilst showing that appeals to experimental methods and findings are required to instruct philosophers on when to trust their intuitions in inquiry. In the next section, I provide an overview of these main positions in recent debates surrounding the experimentalist challenge. This will prove useful to set out the state of play in such debates and to outline the arguments I develop in each chapter of this thesis.

0.4 Examining the experimentalist challenge

0.4.1 The first premise

The first premise of the experimentalist challenge states that intuitions play a central role in philosophy. According to this proposal, philosophical theories are measured (at least in part) by how well they accommodate for our philosophical intuitions. As such, intuitions are commonly regarded to carry significant weight in philosophy, often shaping important philosophical controversies. Typically, this is taken to show that intuitions are *evidence* in philosophy—and, as such, they provide *evidential support* for philosophical positions and claims.

This popular conception of the methodology of philosophy has faced significant criticsm in recent years. Breaking with orthodoxy, a growing number of philosophers contend that intuitions do not play an evidential role in philosophy (Williamson, 2007; Earlenbaugh and Molyneux, 2009; Cappelen, 2012; Ichikawa and Jarvis, 2013; Deutsch, 2015). More forcefully, they argue that it is not at all clear that intuitions *can* provide evidential support for philosophical positions and claims. Thus, they argue that even if philosophers *purported* to use intuitions in an evidential role, intuitions would fail to *be* evidence for philosophical positions and claims.

Furthermore, some Defenders have argued that if philosophers do not rely on intuitions as evidence, then the experimentalist challenge fails at the outset. For, as discussed above, the idea that intuitions play a central role in philosophy is generally interpreted as the claim that they play an *evidential* role. In this sense, some Defenders claim that denying the latter claim is tantamount to denying that they play a central role in philosophy. As a result, empirical findings about intuition turn out to be irrelevant for an assessment of the methodology of philosophy (see, e.g., Cappelen, 2012; Deutsch, 2015).

There are many avenues of response available to Critics to push back against this line of reply to the experimentalist challenge. One of the most promising is to find faults with arguments that seek to establish that philosophers do not rely on intuitions in philosophy. Now, there is no shortage of ammunition for such a reply: a quick survey of the relevant literature reveals discontent with close to every aspect of such arguments (Chalmers, 2014; Weatherson, 2014; Weinberg, 2014; Nado, 2016c; Colaço and Machery, 2017; Brown, 2017). In effect, these objections aim to show that philosophers do indeed rely on intuitions as evidence—and that, as such, that intuitions do play a central role in philosophy.

In chapter 1, I develop an alternative (under explored) line of response to the above objection to the experimentalist challenge. I will argue that even if we accept that philosophers do not rely on intuitions as evidence, this does not rule out that they can play some *other* role of epistemic significance in philosophy. Building on this proposal, I argue that, *pace* some Defenders, empirical findings about intuitions are indeed relevant for assessments of the methodology of philosophy.

0.4.2 The second premise

Defenders have also objected to the second premise of the experimentalist challenge—that is, the claim that experimental findings reveal that people's intuitions are sensitive to a host of epistemically irrelevant factors. The crux of this line of reply is the idea that empirical studies about intuitions display glaring methodological shortcomings. Building on this proposal, Defenders argue for two claims. First, that empirical studies about intuitions are in need of considerable refinements before they can be put to use in debates about the methodology of philosophy. Second, that contrary to what some studies in experimental philosophy suggest, we do not have adequate empirical evidence to conclude that people's intuitions are prone to deleterious effects.

We can here distinguish two ways of developing these claims: piecemeal and general. On piecemeal replies, Defenders aim to find methodological flaws in specific studies in order to show why their conclusions are unwarranted. These piecemeal replies are thus only effective at raising objections about *specific* findings from experimental philosophy—leaving the experimentalist challenge mostly untouched.[8] General replies, on the other hand, seek to raise difficulties about a great number of studies. As such, they raise doubts about the experimentalist challenge overall. In this thesis, I will focus only on general replies and will leave piecemeal responses aside. In what follows, I detail the two most prominent approaches to developing such general lines of reply.

0.4.2.1 Not the right object of study

The first general line of reply to the second premise of the experimentalist challenge claims that extant studies have altogether failed to provide evidence of the workings of intuitions. This proposal begins with the observation that studies in experimental philosophy operate with a minimalist conception of intuition, whereby they are simply the 'judgements people give when presented with actual or hypothetical

[8] Unless, of course, if taken together they amass sufficient evidence to doubt the majority of findings in experimental philosophy.

philosophically-relevant cases.' However, as previously discussed, there is significant disagreement about what counts as a genuine intuition. And, as previously discussed (sec. 0.1.2), many have contested minimalist conceptions of intuition of the sort that experimental philosophers adopt in their studies. Proponents of these more restrictive accounts of intuition thus claim that findings from such studies might be reflecting evidence of the workings of other mental states *besides* intuitions.

Kirk Ludwig (2007) endorses a version of this line of reply. As previously discussed, Ludwig regards intuition to be a kind of judgement that stems solely from on one's conceptual competence. Building on this proposal, Ludwig argues that

> responses to surveys about scenarios used in thought experiments are not *ipso facto* intuitions, that is, they are not *ipso facto* judgement which express solely the subject's competence in the deployment of the concepts involved in them in response to the scenario [...] The task when presented with responses which we know are not (at least all) intuitions is to try to factor out the contribution of competencies from the other factors. This requires an understanding of what the various factors are that may influence responses and enough information about each subject to be able to say with some confidence what factors are at work. It is clear that in the circumstances in which these surveys are conducted we do not have this kind of information.
>
> (ibid., pp. 144-155)

On a very similar note, Kauppinen argues that

> the actual studies conducted so far have failed to rule out competence failures, performance failures, and the potential influence of pragmatic factors, and as such do not yield the sort of results that could support or raise doubts about philosophical appeals to conceptual intuitions.
>
> (Kauppinen, 2007, p. 105)

In effect, both Ludwig and Kauppinen are suggesting that intuitions are the product of careful conceptually rigorous reflection on the matter under evaluation. Furthermore, they contend that experimental studies cannot guarantee that the responses participants give to philosophical cases are judgement of this sort. This then raises the question of whether people's responses in these experimental studies might be reflecting a host factors other than people's conceptual competence regarding the matters under evaluation—e.g., pragmatic distortions, and performance failures. As a result, there seems to be reasons to think that findings from extant experimental studies are unfit for use in assessments of the methodology of philosophy. Call this the 'conceptual competence rejoinder'.

Phenomenalists endorse a similar line of reply to the experimentalist challenge. As previously discussed, the central tenet of phenomenalism is that intuitions have a particular phenomenology—one that makes their contents seem true to a subject. Phenomenalists contend that since studies in experimental philosophy

have altogether failed to control for this putatively *characteristic* phenomenal character of intuition, they are then likely to conflate intuitions with a variety of other mental states. Thus, experimental studies fail to show that *intuitions*—rather than a variety of other mental states—are prone to the deleterious effects identified by studies in experimental philosophy (Bengson, 2013; Chudnoff, 2013, pp. 107-113). As such, they propose that there is insufficient evidence at this point to conclude that the experimentalist challenge is well-motivated. Call this the 'phenomenalist rejoinder'.

Some experimental philosophers have recently responded to the conceptual competence and phenomenalist rejoinders. For instance, Colaço and colleagues (2016) conducted a series of experiments to examine the viability of the conceptual competence rejoinder. To test this proposal, they investigated whether people's intuitions were any different when prompted to think carefully and slowly about the philosophical cases under evaluation. The idea guiding this approach is that such techniques are likely to mitigate pragmatic confusions, and to avoid that participants issue mere hasty responses that could hinder their competence in accurately applying the concepts relevant to the case under evaluation. Colaço et al. reported, however, that people's intuitions were just as likely to vary in problematic ways after such careful reflection. As such, these findings reveal problematic variation even in these cases where people's intuitions are *likely* to reflect their conceptual competence.[9] And Alexander and Weinberg (2014) have recently argued against the phenomenalist rejoinder to the experimentalist challenge. Their main contention is that there is no indication that intuitions (as phenomenalists define them) are any more reliable than other mental states. As such, they claim that the burden of proof is on the phenomenalist to show that the mental states they define as intuitions are not prone to the problematic effects uncovered by studies in experimental philosophy.

I am sympathetic to these objections to both the conceptual competence rejoinder and the phenomenalist rejoinder. However, I find that the reply Alexander and Weinberg advance against the phenomenalist rejoinder leaves the debate in a standstill (as they merely push the burden of proof back on the phenomenalist). In chapter 2, I take on a different tack to responding to the phenomenalist rejoinder. I argue that recent research in human metacognition has studied a class of psychological states that share many of the features that phenomenalists take as characteristic of intuitions. I then survey empirical evidence showing that these psychological states are prone to strikingly similar patterns of variation as those uncovered in studies in experimental philosophy. Building on these considerations, I propose that there are good reasons to think that the phenomenalist rejoinder fails. Finally, I show how experimental research on human metacognition helps to defeat two of the most prominent articulations of the radical version of the experimentalist challenge: viz., that from Unreliability, and from Hopelesness. These arguments will demonstrate how empirical findings on metacognition can help philosophers make better use of intuitions for the purposes of philosophical inquiry.

[9] A proponent of the conceptual competence rejoinder could raise doubts that enforcing participants to engage in careful reflection suffices to show that their evaluations stem from their conceptual competence. However, I agree with Colaço et al., (2016, p. 9) that these kinds of concerns come awfully close to merely *stipulating* that experimental studies are tracking the wrong object of study.

0.4.2.2 Not the right people

Another general reply to the second premise of the experimentalist challenge claims that empirical studies have investigated the intuitions of the wrong people. This line of reply starts from two observations. First, that the relevant empirical studies take non-philosophers to compose their participant pools. Second, that these studies take their findings of problematic influences on the intuitions of non-philosophers to be illustrative of the deficiencies afflicting the intuitions of philosophers as well. However, some Defenders suggest that there are good reasons to deny this assumption. After all, philosophers have years of experience with thinking deeply about philosophical issues, and this presumably shields them from the kinds of mistakes non-philosophers make when evaluating philosophically-relevant cases. This effectively amounts to the claim that philosophers are experts in a way that vindicates their practice of appealing to intuitions in philosophical inquiry. This line of reply has come to be known as the 'expertise defence'.[10]

Proponents of the experimentalist challenge have raised two central objections to the expertise defence. The first relies on findings from research on the nature of expertise, which suggest that rigorous training in an activity is not always conducive to the development of improved performance on that task. These findings underscore that we cannot simply assume that philosophical training will make one better at evaluating philosophically-relevant cases (Weinberg et al., 2010). A second prominent objection comes from recent experimental studies on the intuitions of professional philosophers. Surprisingly, these studies report that the intuitions of philosophers are prone to similar deleterious effects of intra-personal variation as those found in the intuitions of non-philosophers (Schwitzgebel and Cushman, 2012; Tobia, Buckwalter, and Stich, 2013; Schwitzgebel and Cushman, 2015). Building on these sorts of considerations, many experimental philosophers have suggested that there are decisive reason to conclude the expertise defence fails (see, e.g., Machery, 2015; Mizrahi, 2015).

In chapter 3, I argue that the recent tendency in the literature to either defend or refute the expertise defence *simpliciter* is misguided. I contend that both extreme optimism and pessimism about the expertise defence fail to account for crucial differences amongst different *types* of philosophical expertise. I explain that attending to these subtle distinctions reveals that we must turn to piecemeal examinations aimed at testing whether philosophical training shields philosophers from the deleterious effects of *specific* biases on their intuitions. The central upshot of these considerations is that experimental methods can help philosophers make better informed assessments of when to trust their (expert) intuitions, and when to refrain from appealing to intuitions in philosophical inquiry.

0.4.3 The third premise

The third premise of the experimentalist challenge states that if an intuition is sensitive to an epistemically irrelevant factor, then it is unsuitable for use in philosophical inquiry. Perhaps the strongest motivation for this claim arises from focusing on findings of cognitive diversity—that is, findings that people's intuitions vary with respect to their demographic features. To explain, consider the following case:

[10] Proponents of the expertise defence include Kauppinen, (2007), Ludwig, (2007), and Williamson, (2011).

> A philosopher S relies on her intuitive verdict about a given philosophical thought-experiment T to develop her theory of moral responsibility. Experimental studies show, however, that people from cultures other than those of S have a different intuition when presented with T. When challenged on why she has neglected the diverging judgements of people from these other cultures, S just affirms that her own judgements about what obtains in the thought-experiment are *self-evidently* and *obviously* correct. She contends that it is thereby clear that her theory of moral responsibility is correct.

Presumably, there is something amiss with S's attitude here. In particular, it seems clear that it would be epistemically dubious for S to think that her intuition is any better standing than of those from people from other cultures. After all, there do not seem to be good reasons to suppose that belonging to her specific demographic group puts her in a better epistemic standing with respect to matters of morality. In a forceful passage that is representative of this line of thought, Stich suggests that

> unless one is inclined toward chauvinism or xenophobia in matters epistemic, it is hard to see why one would much care that [an intuition] one was thinking of invoking (or renouncing) accords with the set of evaluative notions that prevail in the society into which one happened to be born.
>
> (Stich, 1990, p. 94)

Similarly, Nichols and colleagues remark the following in response to findings purporting to show cultural variation in intuitions about knowledge:

> Without some reason to think that what white, western, high [socio-economic status] philosophers call "knowledge" is any more valuable, desirable or useful than any of the other commodities that other groups call "knowledge" it is hard to see why we should care if we can't have it.
>
> (Nichols, Stich, and Weinberg, 2003, p. 94)

The general idea running through these passages is that cognitive diversity about any thought-experiment T provides *decisive* reasons to think that intuitive judgements about T should pull no weight in philosophical discussions. Likewise, some have argued that a similar verdict applies to intuitions for which evidence of intra-personal variation has been found.

Some Defenders have objected to these strong claims. Perhaps the most prominent of these replies is that by Ernest Sosa. This line of response builds from the run-of-the-mill observation that not all divergences in judgement amount to genuine disagreements. To illustrate, consider the case where two people might appear to disagree over the truth of "There is a bank in town". Suppose, however, that one of them uses "bank" to mean a ground near a river, while the other uses this term to mean the financial institution. In

this case, there is no genuine disagreement between them: rather, they are merely using the word "bank" to express different concepts. Sosa (2009) suggests that findings of cognitive diversity may be uncovering cases of just this sort—that is, instances in which two people are not genuinely disagreeing and are instead just using a same word to express different concepts (and are thus effectively "talking past each other"). Sosa appeals to analogous considerations when responding to the kind of worry expressed by Nichols et al. (2003) in the passage above. He states:

> The fact that we value one commodity, called 'knowledge' or 'justification' among us, is no obstacle to our also valuing a different commodity, valued by some other community under that same label. And it is also compatible with our learning to value that second commodity once we are brought to understand it, even if we previously had no opinion on the matter.
>
> (Sosa, 2009, p. 109)

Although I am somewhat sympathetic to Sosa's claims, I also recognise the pull of Stich's response to them:

> Norms of valuing do play a role in traditional epistemological debates, but they are not the only sorts of norms that epistemologists have considered. [. . .] Goldman insists, quite correctly, that justification rules (or "J-rules") play a central role in both classical and contemporary epistemology, and J-rules specify norms of permissibility not norms of valuing. They "permit or prohibit beliefs, directly or indirectly, as a function of some states, relations, or processes of the cognizer" (Goldman 1986: 60). When we focus on these rules, the sort of pluralism that Sosa suggests is much harder to sustain. If a rule, like the one cited a few paragraphs back, says that ceteris paribus we ought to hold a belief if it is an instance of knowledge, and if 'knowledge' is interpreted in different ways by members of different groups, then Sosa's pluralism leads to inconsistency. There will be some beliefs which we ought to believe on one interpretation of 'knowledge' but not on the other. Moreover, even in the case of norms of valuing, Sosa's pluralism can lead to problems. Sosa is surely right to claim that someone who values owning money banks can also value owning river banks. But if there is one of each on offer and the person's resources are limited, she will have to make a choice.
>
> (Stich, 2009, p. 235)

I will refrain from defending a position in this debate between Sosa and Stich. Instead, in chapter 4 I take on a different tack to show why the third premise of the experimentalist challenge does not obtain. I outline the motivations for this claim, and argue that they rest on a defence of controversial assumptions. Building on these considerations, I argue that intuitions found to be demography-specific *need not* be deemed unsuitable for use in philosophical inquiry. The central upshot of these arguments is that, *pace* Critics, amassing a large body of evidence of cognitive diversity will be insufficient to motivate a radical version of the experimentalist challenge.

0.4.4 The conclusion

The conclusion of the experimentalist challenge is that traditional methods of philosophy stand in need of revisions. As previously stated, proponents of the experimentalist challenge disagree on how to flesh out this claim. Critics argue for the forceful proposal according to which empirical findings show that intuitions do not provide warrant for philosophical positions and claims. As such, they contend that philosophers should revise their methods so as to rule out appeals to intuition in philosophical inquiry. Moderates deny this strong conclusion. On their view, intuitions *can* provide warrant for philosophical positions and claims; however, they contend that traditional methods should incorporate empirical approaches and findings in order to identify and weed out problematic intuitions.

Throughout this thesis, I provide a sustained defence of a moderate version of the experimentalist challenge. In support of this view, I argue against the positions of both Defenders and Critics. These arguments effectively show that neither complete optimism nor pessimism about the use of intuitions in philosophy are justified. Furthermore, they detail how empirical findings from cognitive psychology can help philosophers make better informed assessments of when to trust their intuitions, and when to refrain from appealing to them in inquiry. My arguments for this moderate version of the experimentalist challenge position do not, however, specify why intuitions are poised to provide warrant for philosophical positions and claims. In chapter 5, I address this issue. I explore and develop the proposal that intuitions provide interrogative understanding of their contents. According to this view, a *veridical* intuition that p enables one to understand why p. My approach to develop this proposal is two-fold. First, I argue for a novel contrastive view of interrogative understanding—i.e., the view that to understand why p is to understand why p rather than some alternative, q. I then show how focusing on the contrastive nature of interrogative understanding provides insight into the epistemic state intuitions enable with respect to their contents.

Chapter One

Denying Intuition Denial

In this chapter, I consider and reject a reply to the experimentalist challenge due to Herman Cappelen and Max Deutsch. This line of reply builds on the (increasingly popular) view that intuitions do not play an evidential role in philosophy. I dub this view 'Intuition Denial' (henceforth ID).[1] Both Cappelen and Deutsch argue that if their arguments for ID are correct, then the following claim naturally follows:

Irrelevance Claim: Experimental findings about intuitions are *irrelevant* for assessment of the epistemology of philosophical practices.

Cappelen and Deutsch argue further that, given ID and that this implies the Irrelevance Claim, we should conclude that the experimentalist challenge fails.

In this chapter, I argue that even under the most charitable reading of their arguments, Cappelen and Deutsch fail to motivate the Irrelevance Claim. In the first part of the chapter, I explain how Cappelen and Deutsch purport to motivate the Irrelevance Claim (sec. 1.1) and argue that this proposal does not succeed (sec. 1.2). This discussion will demonstrate that if Cappelen and Deutsch are to successfully motivate the Irrelevance Claim, their arguments must establish that:

Orthogonality Claim: Intuitions play no epistemically significant role *whatsoever* in philosophy.

In the second part of the chapter, I survey each of the three arguments Cappelen and Deutsch advance for ID and evaluate if they could plausibly be taken to motivate the Orthogonality Claim (secs. 1.3, 1.4, and 1.5). I argue that they cannot.

[1] This terminology is adapted from the one used by Jennifer Nado (2016c)

1.1 Motivating the Irrelevance Claim

1.1.1 Intuition Denial

As discussed in the outset of this thesis, intuitions are widely regarded to play a central role in philosophy. On this view, philosophical theories are measured (at least in part) by whether or not they are consistent with certain philosophically-relevant intuitions. Typically, this is taken to show that intuitions provide *evidential support* for philosophical positions and claims. However, this popular conception of the epistemology of philosophy has come under sustained attack in recent years. Breaking with orthodoxy, some philosophers have argued that intuitions do *not* play an evidential role in philosophy (Lycan, 1986; Williamson, 2007; Cappelen, 2012; Deutsch, 2009). I have denominated this view 'ID'.

Proponents of ID agree on the basic tenet of this view. However, their arguments for this position differ in significant ways. For the purposes of this chapter, it will be useful to clarify some of these differences as they prove important for arguments I develop in the subsequent sections.

Williamson develops one of the most influential arguments for ID (Williamson, 2007). His claims for this position are set against the backdrop of what he takes to be a pronounced tendency amongst philosophers to 'psychologise philosophical evidence'. He explains that contemporary philosophers

> think that, in philosophy, ultimately our evidence consists only of intuitions (to use their term for the sake of argument). Under pressure, they take that to mean not that our evidence consists of the mainly non-psychological putative facts which are the contents of those intuitions, but that it consists of the psychological facts to the effect that we have intuitions with those contents, true or false. On such a view, our evidence in philosophy amounts only to psychological facts about ourselves.
>
> (ibid., p. 235)

Williamson argues that this tendency to psychologise philosophical evidence is problematic, as philosophical facts are non-psychological. According to this view, the answer to philosophical questions do not in any way turn on what psychological states philosophers have, or tend to have, when thinking about those issues. Given this view, Williamson argues that by conceiving of our philosophical evidence in psychological terms, we risk divorcing the evidence from the philosophical facts which they are supposed to be evidence for. For, as he states, "psychological evidence has no obvious bearing on many philosophical issues", which concern non-psychological matters (ibid., p. 234).[2] So, Williamson concludes, intuitions do not play an evidential role in philosophy because they *are not* evidence in philosophy. As such, although philosophers *purport* to use intuitions as evidence, their appeals to intuitions fail to provide evidential support to their positions and claims.

[2] Ichikawa (2016) develops an argument that is very similar to this one for ID.

Both Deutsch and Cappelen sympathise with Williamson's suggestion that the psychologisation of philosophical evidence is misguided.[3] However, their arguments for ID take on a rather different tack. Whereas Williamson claims that philosophers *purport* to use intuitions in an evidential role, Deutsch and Cappelen argue, instead, that philosophers make no appeal to intuitions as evidence (Cappelen, 2012; Deutsch, 2015). Their arguments for this radical proposal rely on an innovative exegetical approach: from a close and detailed examination of prominent philosophical texts, they seek to investigate if philosophers appeal to facts about intuitions to provide evidential support for their positions and claims. In effect, this investigation boils down to a project of identifying and cataloguing instances in which philosophers make claims of the sort '*p*, because it seems true that *p*'—which they take as obvious signs of intuitions being used in an evidential role. From such a detailed examination, both Deutsch and Cappelen conclude that philosophers do not purport to use intuitions as evidence after all. As a result, they conclude, it is false that intuitions play an evidential role in philosophy.

In sum, there are two rather different proposals running through these arguments for ID. Williamson's formulation of this view centres on the issue of whether intuitions *can be* evidence for philosophical positions and claims. His suggestion is that they cannot. So, he claims that although philosophers purport to use intuitions as evidence, intuitions simply are not evidence in philosophy. Cappelen and Deutsch, on the other hand, argue that philosophers do not even appeal to intuitions as evidence for their claims in the first place. So, *pace* Williamson, philosophers do not even *purport* to rely on intuitions as evidence for their positions and claims, and so it is false that intuitions play an evidential role in philosophy.

1.1.2 Intuition Denial and the experimentalist challenge

ID has putatively damning implications for the experimentalist challenge. To begin explaining, recall that the experimentalist challenge rests on the assumption that intuitions play a central role in traditional methods of philosophy. Furthermore, as explained above, this claim is typically interpreted as the proposal that philosophers rely on intuitions as *evidence* in philosophical inquiry. But, if arguments for ID are correct, then philosophers do not rely on intuitions as evidence after all. Thus, ID suggests that the experimentalist challenge rests on a mistaken conception of the methodology of philosophy (and so it fails outright).

Cappelen provides an exceptionally clear formulation of this line of reply to the experimentalist challenge. He states that, given his arguments for ID, it then becomes obvious that

> studies of intuitions people have about thought experiments have *no direct relevance* for philosophical arguments or theorizing. [...] In short: If philosophers don't rely on intuitions [as evidence], then the project of checking people's intuitions is *philosophically pointless*.
>
> (Cappelen, 2012, pp. 221-222) [4]

[3] See Deutsch, (2015) and Cappelen, (2012, p. 204)

[4] Italics are mine

And on a similar note, Deutsch claims that

> it really is simply a myth that philosophers employ a method, in reasoning about thought experiments and cases, whereby they make evidential appeals to intuitions. [...] As a result, the data collected by [experimental philosophers] in intuition surveys are *mostly* irrelevant to analytic philosophical method.
>
> (Deutsch, 2015, p. xvi) [5]

It might seem that Deutsch hedges commitment to the strong claim that ID threatens the experimentalist challenge when he states that empirical findings are "*mostly* irrelevant" to philosophical method. However, he clarifies that what he means is that findings about intuitions *might* be "relevant to philosophy in a broadly *ethical* way'', as they could be "relevant to how we should treat others and how, more fundamentally, we should understand the social practices of different groups of people" (ibid., p. 160). In this sense, Deutsch expresses hope that his arguments for ID will bring the focus to *the* useful facet of empirical findings about intuition—viz, these ethical considerations (ibid., p. 161).

In effect, these passages claim that if ID is true, then findings about intuitions are irrelevant to assess the epistemology of philosophy (I have dubbed this the Irrelevance Claim). That is, they take the following conditional to hold:

ID/IC : If ID is true, then the Irrelevance claim is true.

Now, it is quite clear that the Irrelevance Claim spells trouble for the experimentalist challenge. After all, if findings about intuitions have no relevance for assessments of the epistemology of philosophy, then these findings fail to motivate any concerns about the methodology of philosophy.[6]

There are many ways a Critic might seek to push back against this objection to the experimentalist challenge. Perhaps the most promising avenue of response is to argue that extant attempts to establish ID fail. Now, there is no shortage of ammunition for such a reply. For instance, a quick survey of the recent literature on this topic reveals discontent with close to *every single* aspect of Cappelen and Deutsch's claims for ID (Chalmers, 2014; Weatherson, 2014; Weinberg, 2014; Nado, 2016c; Colaço and Machery, 2017). Williamson's version of ID has also faced significant criticism, on the basis that it rests on a set of highly controversial epistemological assumptions (Brown, 2017).

Proponents of ID have attempted to respond to some of these objections (Cappelen, 2014; Deutsch, 2015). This has in turn sparked renewed interest in ID, and on the Irrelevance Claim. For the purposes of this chapter I will sidestep all such debates. Instead, I will grant (for the sake of argument) that *all* readings of ID are correct. As such, I will assume that proponents of ID can adequately respond to the above critiques

[5] Italics are mine.

[6] Not all proponents of ID accept ID/IC (see, e.g., Ichikawa, 2014).

to ID. However, in the next section I argue that even if we grant ID, this falls short of establishing the Irrelevance Claim. As a first step to developing these claims, I begin the next section by arguing that ID/IC is false. I then explain why this shows that Williamson's version of ID is *compatible* with the Irrelevance Claim. But, as I explain later in the section, denying ID/IC is not tantamount to showing that Cappelen and Deutsch outright fail to motivate the Irrelevance Claim, as there is at least one line of reply available to them that warrants further investigation. I examine this objection in more detail in the subsequent sections.

1.2 Irrelevance Claim and ID/IC

1.2.1 Against ID/IC

In this section, I argue that ID/IC is false: the Irrelevance Claim does not immediately follow from ID. Key to this proposal is an appreciation of the fact that questions about *evidence* do not exhaust all and every epistemically significant dimension of our practices of inquiry.[7] To begin explaining, consider the following scenario:

> **Errant**: Psychological studies have discovered a pernicious cognitive bias that affects how people respond to a body of evidence E. This cognitive bias manifests in the tendency to neglect certain propositions contained in E, which then makes p seem true. As q clearly follows from p, people prone to this bias tend to infer q when presented with E.
>
> There is simply no agreed upon answer as to what is the normatively correct response on whether p given E.

There are two noteworthy observations to make about **Errant**. The first is that the bias described in this case does not reasonably play an *evidential* role in inquiry about q. Facts about this bias do not, after all, speak in favour of either q or *not-q*. At most, they show that people are likely to respond in certain ways on whether p when presented with E; but, this is neither here nor there with respect to whether q is true or not. Second, although the bias described in **Errant** does not play an evidential role in inquiry about q, it nevertheless does play *an* epistemically significant role in this line of inquiry.[8] After all, facts about this bias are relevant for how we should form beliefs when inquiring into whether q. For instance, learning that one's belief that q is caused by the bias described above would compel one to revise that belief. And, in this sense, this bias can reasonably be said to be epistemically significant when inquiring into whether q. Thus, **Errant** describes one instance in which an object is epistemically relevant for a line of inquiry, even if it does not play an evidential role within it.

With these considerations in mind, let us now turn to ID/IC. What is important to note is that the above observations regarding **Errant** suggest that ID is compatible with a denial of the Irrelevance Claim. To

[7] The views I advance in this section draw heavily on the ideas developed by Jonathan Ichikawa (2016) and Jennifer Nado (2016c)

[8] The arguments I develop here rely on the widespread assumption that the goal of inquiry is factive.

begin explaining, note that this discussion reveals that questions about evidential support do not exhaust all and every *epistemically* significant dimension of our practices of inquiry. For example, it is clear that facts about the cognitive bias described above do not provide evidence for the proposition q. Nevertheless, it seems clear that facts about this bias can still play a role of epistemic significance in inquiry into whether q. This suggests, in turn, that even if we accept ID, and that intuitions do not play an evidential role in philosophy, this does not suffice to show that intuitions play no epistemically significant role whatsoever. For, even if proponents of ID are right to think that philosophers do not rely on intuitions as *evidence*, this would still leave it open that they rely on intuitions in *a* role of epistemic significance.

What would this non-evidential, epistemically significant role of intuitions be? One proposal is that, like the bias described in **Errant**, facts about what claims philosophers find intuitive have significant bearing on how philosophers respond to the evidence available to them. Thus, even if an intuition is neither here nor there with respect to what that evidence supports, it might nevertheless influence what claims philosophers come to accept when presented with that evidence. Moreover, there are good reasons to think intuitions play an epistemically significant role in philosophy. After all, philosophers often report believing claims *because* they find them intuitive—e.g., because it seems obvious to them that the subject in the Gettier case does not know, or that it is immoral to push the man to stop the run-away trolley. Surely, this does not *entail* that philosophers only believe those propositions once they are intuited. But, this provides strong reasons to believe that because philosophers are more likely to believe intuitive propositions, that these intuitions influence what a body of evidence supports—and as such play *some* role in belief formation.[9]

In sum, these considerations suggest that even if ID is true, this leaves it open that intuitions can play some role of epistemic significance in philosophy. Now, note that this conclusion effectively shows that ID/IC fails. For, if intuitions do play such a (non-evidential) epistemically significant role in philosophy, then it is clear that findings about intuitions are relevant to assess the methodology of philosophy—that is, the Irrelevance Claim fails. This then indicates that the conditional claim expressed in ID/IC does not obtain: the Irrelevance claim does not immediately follow from ID.

1.2.2 Replies

To a first approximation, the arguments in the previous section show that attempts to establish the Irrelevance Claim via arguments for ID do not succeed. For, as explained above, there is good reason to think that ID is entirely consistent with a denial of the Irrelevance Claim. This in turn suggests that ID/IC is false. As such, it would seem obvious that extant arguments for ID fail to motivate the Irrelevance Claim. In this section, I argue that whilst this verdict applies to Williamson's arguments for ID, there is one line of reply available to Deutsch and Cappelen that warrants further investigation. To develop these claims, I will consider the two most promising lines of reply to the arguments I developed against ID/IC.

[9] Curiously, Deutsch (2009, p. 451) even accepts that intuitions could have such a role in philosophical inquiry. However, he contends that this would amount to nothing more than a *causal* function, with no bearing whatsoever on matters of epistemic justification.

The objections I developed against ID/IC relied on the following conditional: if intuitions play a (non-evidential) epistemically significant role in philosophy, then findings about intuitions are relevant to assess the methodology of philosophy. One clear way to deny my objections would be to deny the antecedent of this conditional. Here is one way to do so: one could maintain that *all* questions in Epistemology are reducible to matters of evidential support. In effect, this proposal boils down to the claim that there is no other epistemically significant role besides evidential roles. On this view, the ID/IC clearly obtains. For, this would effectively entail that if intuitions do not play an evidential role in philosophy, there would be no epistemically significant role they *could* play. As such, the Irrelevance claim would immediately follow from ID.

The main problem for this proposal is that there does not seem to be any good reason to think that all debates in epistemology are reducible to questions about evidence. To explain, consider first the notion of epistemic blame. The proposal that a subject can be blameworthy for violating an epistemic norm has become standard in a number of central philosophical debates. And importantly, some 'radical externalists' in Epistemology take evaluations of epistemic blame to be separable from questions of evidential support (Williamson, 2013; Littlejohn, forthcoming). In addition, consider cases of beliefs caused by wishful thinking. For instance, suppose that I form a belief that there is chocolate in the kitchen because I want there to be some chocolate there. Now, nobody would claim that wishful thinking of this sort is evidence for the belief that there is chocolate in the kitchen; nevertheless, discovering that this is what caused that belief provides decisive reasons to revise it. In this sense, it is quite clear that wishful thinking is epistemically significant, even though it is evidentially inert. Similar conclusions can be drawn by reflecting on the epistemically significant (albeit non-evidential) role of priors in Bayesian epistemology. Taken together, these considerations provide strong reason to doubt that questions about evidence account for all and every aspect of Epistemology. And, even if this view is tenable, it would be a monumental task to provide reasons in support for why we should consider it a viable position. As proponents of ID have not provided any such support, I see no reason to accept it.

So much for ID/IC. Now, it is noteworthy that even if proponents of ID fail to motivate ID/IC, this does not mean that they thereby fail to motivate the Irrelevance Claim. For, it might be suggested that various of the arguments proponents of ID develop for this claim lend support to the idea that intuitions play no epistemically significant role *whatsoever* in philosophy—I dubbed this the 'Orthogonality Claim'. Indeed, both Cappelen and Deutsch (although not Williamson) at points seem to suggest that their arguments should be taken to motivate this stronger claim. For example, Cappelen explicitly suggests that his arguments for ID show that it is false that "[p]hilosophers rely (*in some epistemically significant way*) on intuitions when they make judgements about cases" (Cappelen, 2012, p. 192).[10] And Deutsch affirms that his arguments for ID show that *all that matters* for an assessment of a philosopher's epistemic standing with respect to her positions and claims is whether the *arguments* she advances for them are correct (Deutsch, 2015). In this sense, Deutsch seems to suggest that considerations about whether those arguments are correct encompass *all and every* epistemically significant dimension of assessment of

[10] Italics are mine

philosophical inquiry—which would then effectively entail that intuitions fail to play any epistemically significant role in philosophy.

Thus, there does seem to be at least some indication that Cappelen and Deutsch think their arguments are meant to motivate the Orthogonality Claim. And notably, the Orthogonality Claim effectively implies the Irrelevance Claim. For, if intuitions are thus epistemically otiose, then empirical findings about intuitions simply cannot be relevant for debates about the epistemology of philosophy. Thus, if Cappelen and Deutsch's arguments for ID could indeed be taken to motivate the Orthogonality Claim, they would thereby vindicate the Irrelevance Claim. For the remainder of this chapter, I assess whether this strategy for motivating the Irrelevance Claim is viable. In the subsequent sections, I survey each of the three arguments Cappelen and Deutsch develop in support of ID and evaluate whether they could reasonably be taken to motivate the Orthogonality Claim. My conclusion is that they cannot. Building on these evaluations, I propose that even under the most charitable reading of their arguments, Cappelen and Deutsch simply fail to motivate the Orthogonality Claim. As such, they fail to establish the Irrelevance Claim, and to refute the experimentalist challenge.

1.3 The Argument from Exegesis

1.3.1 The original argument

The main argument Cappelen and Deutsch advance for ID claims that philosophical texts provide no reason to believe that philosophers use intuitions in an evidential role. To argue for this proposal, Cappelen engages in a detailed textual analysis of ten prominent philosophical thought-experiments—many of which are widely regarded to be paradigm cases of appeal to intuitions in philosophy. The aim of this careful exegetical approach is to conduct what he denominates an 'intuition-diagnostic': a procedure to identify appeals to intuitions in philosophical argumentation. To develop this approach, Cappelen first selects what he takes are commonly-recognised "features of the intuitive"; then, he turns to philosophical texts to count how many times philosophers explicitly appeal to anything with these features. From this close reading, Cappelen concludes that "actual instances of philosophical argumentation hardly ever exhibit signs of relying on intuitions [as evidence]" (Cappelen, 2012, p. 187). Deutsch takes on a very similar exegetical approach. He first selects prominent philosophical texts and counts the amount of times their authors use 'intuition' and other cognate terms. Deutsch's gives a close reading of what are perhaps the two most influential thought-experiments in recent analytical philosophy: namely, Kripke's Gödel case against descriptivist theories of semantic reference, and Gettier's thought-experiments against the reductive analysis of knowledge in terms of justified true beliefs. From this in-depth examination, Deutsch concludes that "close attention to how Gettier and Kripke actually argue fails to reveal any explicit or implicit commitment to using intuitions" (Deutsch, 2015, p. 40).

Cappelen and Deutsch take their intuition-diagnostic procedure to reveal that the authors from the texts examined do not rely on intuitions as evidence. Moreover, both take their textual analyses to warrant broader conclusions about the philosophical discipline. Specifically, they suggest that the patterns they

reveal in these texts are representative of the philosophical discipline as a whole. Thus, they contend that the conclusion they draw about the practices of philosophers from these few writings reveal that philosophers *in general* do not rely on intuitions in philosophical argumentation. This suggests that we should conclude that philosophers do not use intuitions in an evidential role. More schematically:

The Argument from Exegesis (AEx)

P1 Textual analysis shows that philosophers do not make explicit appeals to intuitions in an evidential role.

P2 If philosophers do not make explicit appeals to intuitions in an evidential role, then intuitions do not play an evidential role in philosophy.

C Intuitions do not play an evidential role in philosophy. (ID)

The AEx has attracted a lot of criticism. We can divide these replies in two groups, depending on which premise they reject. Most objections of the AEx target the first premise of this argument. These replies express discontent with what Cappelen and Deutsch regard as the diagnostic features of intuitions—that is, those characteristics they take as distinctive of intuitions, and which they rely on to identify appeals to intuitions in philosophical texts. In particular, many suggest that these features are inadequate to demarcate intuitions from other mental states. Indeed, as discussed in the outset of this thesis, there is heated ongoing dispute about how to best characterise intuitions. And notably, many views on the nature of intuitions are incompatible with those that Deutsch and Cappelen employ in their textual analyses. This suggests that the textual analysis Cappelen and Deutsch provide may simply fail to establish that their authors do not explicitly appeal to intuitions in their writings. After all, if their exegetical approach wrongly defines what intuitions are, then they are prone to neglecting obvious signs of appeals to intuitions in an evidential role in those texts.

Another prominent set of objections to the AEx raise doubts about its second premise. The central claim motivating this objection is the humdrum observation that we often fail to make explicit what is the evidence for the claims we make. For instance, if I assert that the prices of sardines in the supermarket have dropped, I might not state that I believe this because I saw the new prices on the supermarket shelf. Now, just because I fail to make explicit that my perceptual evidence is what grounds my statement that sardines are cheaper on sale, this in no way demonstrates that I did not rely on that perceptual evidence when forming that belief. On a similar note, it seems clear that even if philosophers do not make it explicit in their writings that they rely on intuitions as evidence for their positions and claims, this fails to rule out that they might be doing so after all (Nado, 2016c; Brown, 2017). This then raises doubt about the second premise of AEx.

I am sympathetic with the main idea of the above objections. But, as I have previously indicated, I will accept that Cappelen and Deutsch have successfully established ID. As such, I will assume they can respond to these replies to the AEx. However, I maintain that the AEx still falls short of establishing

the Irrelevance Claim. For, as I argued previously, the only way Cappelen and Deutsch could seek to vindicate the Irrelevance Claim is if their arguments could plausibly be taken to motivate the stronger Orthogonality Claim—i.e., the claim that intuitions play no epistemically significant role whatsoever in philosophy. But, can we reasonably reformulate the AEx as a successful defence of the Orthogonality Claim? For the remainder of this section, I argue we cannot.

1.3.2 The revised argument

There are two ways to read the AEx such that it could purport to motivate the Orthogonality Claim. The first would require a reinterpretation of the second premise of this argument—i.e., the claim that 'If philosophers do not make explicit appeals to intuitions in an evidential role, then intuitions do not play an evidential role in philosophy'. To get the Orthogonality Claim off the ground, we could reinterpret it to read that 'If philosophers do not make explicit appeals to intuitions in an evidential role, then intuitions do not play any epistemically significant role *whatsoever* in philosophy.' Plugging this into the AEx would yield the following argument:

Argument from Exegesis revised (AExr)

P1 Textual analysis suggests that philosophers do not make explicit appeals to intuitions in an evidential role.

P2 If philosophers do not make explicit appeals to intuitions in an evidential role, then intuitions do not play any epistemically significant role in philosophy.

C Intuitions play no epistemically significant role whatsoever in philosophy. (Orthogonality Claim)

This argument is valid. However, there are at least two interrelated reasons to resist the second premise of this revised version of AEx. First, note that the second premise of the AExr virtually suggests that *all that matters* for an evaluation of the epistemology of philosophical inquiry are considerations about evidential support. On this proposal, an analysis of the items that play an evidential role in philosophy suffices to provide an assessment of the methodology of philosophy. But, as I argued in the previous section, this position is controversial. For, this proposal is tantamount to the claim that all debates in Epistemology are reducible to questions about evidential support. And, there are at least a few cases where it seems fairly clear that this proposal would fail (e.g., questions about epistemic blame). So, if we are to take the AEx seriously, Cappelen and Deutsch would need to provide arguments to the contrary. As they provide no such arguments—and since none seem to be forthcoming—I take it there is no reason to endorse the AExr.

A second more promising way of revising the AEx so as to motivate the Orthogonality Claim would be to reinterpret its first premise—i.e., the claim that 'Textual analysis suggests that philosophers do not make explicit appeals to intuitions in an evidential role.' One way to reinterpret this premise would be to suggest that the textual analyses of Cappelen and Deutsch demonstrate, not only that philosophers do

not rely on intuitions in an evidential role, but that they do not rely on intuitions in any epistemically significant role whatsoever. This would then yield the following argument:

Argument from Exegesis revised 2 (AExr 2)

P1 Textual analysis suggests that philosophers do not make explicit appeals to intuitions in any epistemologically significant role.

P2 If many philosophers do not make explicit appeals to intuitions in an epistemically significant role, then there is no reason to think that philosophers ever appeal to intuitions in any *epistemically significant role whatsoever* in philosophical inquiry.

C Intuitions play no epistemically significant role in philosophy. (Orthogonality Claim)

The AExr 2 is also valid. However, I propose we have good reasons to reject its second premise. My objection echoes a reply to the *original* AEx (discussed earlier in this section). On this line of reply, an examination of whether philosophers explicitly mention appeals to intuitions in support of their claims is inadequate to identify whether they rely on intuitions in an evidential role. This reply begins with the humdrum observation that we don't often make explicit reference to what is the *evidence* for the claims we make. I may, e.g., believe that the prices of sardines in the supermarket have dropped because I *saw* their price tags on the shelves. Now, this shows that even if I fail to note my perceptual evidence when I claim that the price of sardines have dropped, this does not show that I did not rely on my perceptual evidence after all. Likewise, we should not take a failure to mention appeals to intuitions in an evidential role in philosophy as a demonstration that intuitions are not used as evidence in philosophy. Now, I submit that this line of reply naturally carries over to account for other roles of epistemic significance, beyond evidential ones. For, just as we do not often make explicit what plays an evidential role in belief formation, we also often fail to note what plays other (non-evidential) roles of epistemic significance in such processes.[11] For example, suppose that a pernicious cognitive bias causes me to neglect certain propositions from a body of evidence E when presented with E, which in turn makes it seem to me that p is true. Now, even if I do not make it explicit that my judgement that p was at least partially due to the influence of this cognitive bias, this does not demonstrate that this bias is epistemically insignificant with respect to my judgement that p. After all, it is undeniable that facts about this bias are of epistemological significance for inquiries into whether p. Similarly, consider the case in which I take a philosophical proposition q to be true because I find it to be intuitive. Now, what if I then come to use q in philosophical theorising without mentioning that I endorse q because I had an intuition that it is true? Does this show that my intuition that q did not play an epistemically significant role in my belief that q? Presumably not. In this sense, examination of whether or not I mention having an intuition that q is inadequate to assess whether I appealed to that intuition in an epistemically significant role in belief-formation about q. This

[11] At this point, the reader might baulk at my suggestion that Deutsch and Cappelen could successfully avoid this line of reply when applied to the original version of the AEx. Here's how: I take it that Deutsch and Cappelen could rely on their endorsement of Williamson's claims that intuitions are not evidence in philosophy.

then suggests that we should resist the second premise of the AExr 2 and that, as a consequence, this revised version of the AEx fails.

In sum, I have discussed the two charitable readings of the AEx in which they purport to motivate the Orthogonality Claim. However, I argued that these revised versions of the AEx are ultimately untenable. The first requires endorsing the controversial position according to which questions about evidence are all that matters to assess the epistemology of philosophical inquiry. But, as pointed out in the previous section (sec. 1.2), there are good reasons to reject this proposal; and so, this first revised version of the AEx fails. The second revised version suggests that the exegetical analysis Cappelen and Deutsch provide demonstrates that philosophers do not rely on intuitions in any epistemically significant role whatsoever. However, I have suggested that this exegetical approach is not suitable for this end, and cannot plausibly vindicate this claim. In conclusion, I submit that the most plausible revised versions of the AEx fails to motivate the Orthogonality Claim. As a result, they also fail to vindicate the Irrelevance Claim—and cannot thus undercut the experimentalist challenge.

1.4 The Argument from Alternative Explanation

1.4.1 The original argument

The central upshot of ID is that, despite what philosophers might think, they do not appeal to intuitions as evidence for their positions and claims. In effect, this amounts to the claim that philosophers are *collectively* mistaken about what practices they employ in the course of philosophical inquiry. Now, it is noteworthy that this is quite a radical proposal. How could so many philosophers be mistaken about the practices central to their profession? This question make salient that ID is questionable lacking a supplementary picture to fill in the 'intuition-shaped hole' left in philosophers self-conception of their discipline.

Cappelen and Deutsch are both sensitive to this concern. Their second main argument for ID addresses this issue by attempting to provide an alternative picture of the epistemology of philosophy. For this aim, Cappelen and Deutsch once more make recourse to an exegetical approach that examines prominent philosophical writings. The purpose of this analysis is to identify that which philosophers do offer as evidence for their views. To explain this proposal, it will be useful to provide one example of this approach. For this aim, I focus on what is perhaps the most detailed formulation of this alternative epistemological picture: namely, Deutsch's reconstruction of Gettier's 10 coin case.

Like most philosophers, Deutsch agrees that the correct verdict on the scenario described in that case is that Smith does indeed have a justified true belief that (e), yet does not *know* that (e). But, against orthodoxy, Deutsch objects that what justifies or provides evidence for this verdict is the fact that people are likely to have an *intuition* that Smith does not know that (e). For, Gettier nowhere makes explicit appeal to what is 'intuitive', or mention that we are supposed to have any particular 'intuition', as evidence for this conclusion. Rather, Deutsch points out, a textual analysis reveals that Getter offers *only* the following argument in support of the verdict that Smith does not know that (e):

> But it is equally clear that Smith does not know that (e) is true; for (e) is true in virtue of the number of coins in Smith's pocket, while Smith does not know how many coins are in Smith's pocket, and bases his belief in (e) on a count of the coins in Jones's pocket, whom he falsely believes to be the man who will get the job.

Deutsch argues further that if Gettier offered only this argument in support of the claim that Smith does not know that (e), then we should read him to regard that argument as *the* evidence for this claim. As such, Deutsch suggests that the correct interpretation of Gettier's writings is that he relies only on the above argument to provide evidential support to the conclusion that Smith does not know that (e).

Cappelen and Deutsch offer a similar verdict about all other philosophical texts they examine. In more detail, they first suggest that textual analyses reveal that the authors of these texts offer only arguments in support of the claims they advance. From this analysis, they suggest that those philosophers are relying on only those arguments as evidence for their positions and claims. Furthermore, Cappelen and Deutsch both suggest that these analyses are representative of the philosophical discipline as a whole. As such, they take an inductive step to then claim that philosophers *in general* rely only on arguments as evidence. Call this the 'Argument from Alternative Explanation' (AAE):

Argument from Alternative Explanation (AAE)

P1 Textual analysis reveals that philosophers offer *only* arguments as evidence for their philosophical positions and claims.

P2 If philosophers offer only arguments as evidence of their positions and claims, then they rely on only those arguments as evidence for their positions and claims.

C Philosophers rely only on arguments as evidence for their philosophical positions and claims. (ID)

The AAE has also attracted a fair amount of criticism. For the most part, these objections have focused on the second premise of this argument. The main thrust of such replies is the claim that, even if philosophers offer only arguments as evidence, this still leaves it open that they rely on intuitions in an evidential role after all (Chalmers, 2014; Weatherson, 2014; Nado, 2016a; Brown, 2017). To begin explaining, consider again the suggestion that a textual analysis reveals that Gettier offers only the above brief argument in support of his verdict on the 10 coin case. Now, even if we grant this much, it does not identify what Gettier took as evidence for the premises of *that* argument. At this point, a natural suggestion would be to say that Gettier took his intuitions to be evidence for those premises. Of course, Cappelen and Deutsch would deny this. But then, this pushes the burden of proof back onto them to show what is the evidence for those premises. In line with the AAE, we could entertain the hypothesis that Gettier just relied on even further arguments in support of such premises. But, it is clear that this in turn invites an infinite regress—with each of these arguments requiring another argument in support of their own premises. Again, a natural suggestion is that intuitions might come in at some point to end this justificatory chain.

Given that the textual analyses that Cappelen and Deutsch offer cannot rule out this possibility, then they fail to show that philosophers do not rely on intuitions in an evidential role.

Deutsch has recently attempted to respond to these objections (Deutsch, 2015, chs. 4 & 5). In a nutshell, his arguments amount to the claim that the challenge of accounting for the ultimate justifier is a problem for all philosophers—and not just for him. And so, he contends that it is not so clear that the burden of proof is on him to show that intuitions do not come in at any point in this justification chain. I will simply assume that this response succeeds. The motivation for this is that as I have indicated in the outset of this chapter, I am here granting ID (just for the sake of argument). However, I submit that even if we grant the AAE, this argument cannot by itself establish the Irrelevance Claim. Again, the reason for this is that the AAE can at most provide support for ID. But, as I have argued, Cappelen and Deutsch must vindicate the Orthogonality Claim if they are to adequately motivate the Irrelevance Claim. So, can we revise the AAE so as to successfully motivate the Orthogonality Claim? For the remainder of this section, I will argue that we cannot.

1.4.2 The revised argument

The AAE easily lends itself to a charitable reading so as to motivate the Orthogonality Claim. The most promising of these readings requires a modification of the first premise of the AAE—that is, the claim that textual analysis reveals that philosophers offer only arguments as evidence for their philosophical positions and claims. On this alternative reading, the textual analyses provided by Cappelen and Deutsch demonstrate, not only that philosophers offer exclusively arguments as *evidence*, but that these arguments are the single source of *epistemic support* they provide for their positions and claims. In other words, this would amount to the claim that all of the epistemic support these views receive come from such arguments. Furthermore, in line with Cappelen and Deutsch's exegetical approach, we can then conclude that philosophers rely only on such arguments as epistemic support for their claims. Now, note that this then plausibly implies that the epistemic standing of those positions and claims hinges *exclusively* on the arguments philosophers advance. As a consequence, it becomes hard to see what epistemically significant role intuitions could play in philosophy. That is:

Argument from Alternative Explanation revised (AAEr)

P1 Textual analysis suggests that philosophers offer only arguments in support of their philosophical positions and claims.

P2 If philosophers offer only arguments in support of their positions and claims, then they rely on only those arguments in support of their positions and claims.

P3 If philosophers rely on only arguments in support for their positions and claims, then the epistemic standing of their positions and claims depends *exclusively* on those arguments.

C Intuitions play no epistemically significant role in philosophy.

I submit there are good reasons to resist both the second and third premises of this argument. Let us start with the second premise—i.e., the claim that 'If philosophers offer only arguments in support of their positions and claims, then they rely only on those arguments as epistemological support for their positions and claims.' The objection I advance to this premise resembles a line of reply I discussed with regards to the AExr (see previous section). Again, the objection is that focusing on what philosophers explicitly offer in support of their claims is inadequate to determine what they rely on. For instance,we often don't make explicit reference to what is the evidence for the claims we make. Now, we can advance a similar line of reply with respect to the second premise of the AAEr. Specifically, we can simply claim that because philosophers offer only arguments in support of their claims, this is insufficient to show that these arguments are all that they rely on as epistemological support for those claims. For instance, even if philosophers fail to mention that they endorse an argument because they find their premises intuitive, this does not mean that they do not rely on those intuitions as epistemological support for those arguments after all.

Let us now turn to the third premise of the AAEr—i.e., the claim that 'If philosophers rely only on arguments as support for their positions and claims, then the epistemic standing of their positions and claims depends exclusively on the arguments they advance.' Now, it is not too hard to see that attempting to account for the epistemology of our philosophical practices exclusively in terms of the arguments philosophers advance invites some bad results. For instance, suppose that Gettier took the above argument to be correct because his astrologist read this off of his zodiac map. Presumably, it is less than rational to accept an argument because one believes in the power of the stars to deliver philosophical truths. In this sense, it seems false to assert that all that matters for an evaluation of Gettier's epistemic standing with respect to the claim that 'Smith does not know that (e)' is that he endorses a correct argument to this effect. Similarly, we can find fault in a student who believes Gettier's argument merely because she heard a professor claim that this argument is true. Presumably, this alone does not suffice to put her in a good epistemic standing with respect to the claim that Smith does not know that (e). For, unless she has a good *reasons* to accept Gettier's argument for this conclusion, she could not be said to be in a sufficiently good position to endorse it. So, to a first approximation, it seems that focusing exclusively on the arguments a philosopher advances for her positions and claims would presumably fail to account for important dimensions of the epistemology of our philosophical positions and claims. After all, it would only seem reasonable to think that the reasons for which either Gettier or anyone else would have to hold these arguments to be true are an important feature of whether they are in a good epistemic standing to endorse.[12]

In light of these two objections, I submit that the AAEr fails. Against its second premise, I argued that an analysis of what philosophers explicitly offer in support of their claim is not reliable indication of what they actually relied on as epistemological support for those claims. Against the third premise, I argued that we cannot provide a satisfactory analysis of the epistemology of a philosopher's positions and claims by focusing exclusively on the arguments she advances for them. As such, I contend that the AAEr fails

[12] See Brown (2017) for a similar argument.

to motivate the Orthogonality Claim. As a consequence, it cannot motivate the Irrelevance Claim.

1.5 The Argument from Gibberish

1.5.1 The original argument

Cappelen advances one further argument for ID. This argument claims that, given the variability in how philosophers use the term 'intuition', it is not clear what this term refers to. As such, he claims that it is not clear how we should interpret the claim that 'philosophers rely on intuitions in an evidential role'. To set up this argument, Cappelen first runs through many applications of the term 'intuition' in ordinary discourse and shows how they diverge from the way in which philosophers talk about intuitions (36–44). He then suggests that the mismatch between ordinary and philosophical uses of the term entails that philosophers *must* be using 'intuition' in a technical sense. That is, he suggests that the philosophical community agrees that 'intuition' picks out a given referent, but that is not the one denoted by ordinary uses of that term. However, Cappelen points out that disagreement between philosophers about the nature of intuitions is rampant. As a quick survey of the literature on the topic reveals, philosophers hold a diversity of incommensurable views about the nature of intuitions, which have been described as

1. beliefs or judgements (Lewis, 1983; Williamson, 2007)

2. inclinations to believe or judge (Inwagen, 1997a; Sosa, 2007b)

3. a *sui generis* class of mental states (Bealer, 2000)

All of which have been proposed to have a distinctive

- phenomenology (Koksvik, 2011; Chudnoff, 2013; Bengson, 2015b)

- content (Bealer, 2000; Sosa, 2007b)

- etiology (Ludwig, 2007; Grundmann, 2010)

Noting this diversity of views, Cappelen contends that philosophers cannot be using 'intuition' in any technical sense. As a result, Cappelen concludes that philosophers' talk of intuition is semantically defective: it fails to secure a specific semantic content, making it "strictly speaking, meaningless" (Cappelen, 2012, p. 59). Otherwise put, philosophers' talk of 'intuition' fails to denote any particular mental state or event.[13] Building on this suggestion, Cappelen seems[14] to argue further that talk of 'intuitions' fails to pick out any specific thing that *could* be used as evidence in philosophical inquiry. Call this the Argument from Gibberish (AG). More schematically:

[13] Notably, Cappelen at one point claims that his arguments for ID do not rely on the idea that philosophers' talk of intuitions is defective (Cappelen, 2012, p. 60). However, he does seem to rely on this idea later when developing the AG (ibid., p. 63).

[14] I qualify this statement with 'seems' because it is not at all clear why Cappelen takes the defectiveness of intuition talk to motivate ID. And notably, many have expressed concerns that it cannot (see, e.g. Anna-Sara, 2013).

Argument from Gibberish (AG)

P1 Philosophers' talk of 'intuitions' is defective: it fails to denote any particular mental state or event.

P2 If philosophers' talk of 'intuitions' is defective, then it fails to pick out something that could play an evidential role in philosophy.

C Philosophers do not rely on intuitions as evidence for their philosophical positions and claims. (ID)

The AG has faced significant criticism. For the most part, these critiques take issue with the first premise of this argument. According to this line of reply, even if philosophers' talk of 'intuition' neither conforms to ordinary language or to a technical use of the term, it still successfully refers to *a* particular mental state or event (Weatherson, 2014; Bengson, 2014). For instance, John Bengson (ibid.) suggests that philosophers adopt a *discriminative* use of 'intuitions' that "while rooted in natural language, is more discriminating (i.e., it has a narrower extension) than what is found in casual ordinary discourse." Bengson argues that this discriminative use is revealed in how philosophers use the word 'seems' to "describe how one is intellectually appeared-to, or, better, how things are presented to one as being, when one reflects on them" (ibid., p. 562), and to pick out an "intellectual state or event with some epistemic significance" (ibid., p. 564). Michael Devitt (2015) also objects to the first premise of the AG, for he contends that philosophers' talk of intuition (neatly) picks out a kind of unreflective, immediate judgement without reasoning or inference.

I propose, however, that the first premise of the AG is indeed correct. My chief motivation for this claim comes from experimental and theoretical developments in psychology on the study of the cognitive underpinnings of what are generally termed to be 'intuitive' responses. What these studies demonstrate is that our intuitions are generated by very different sorts of cognitive processes (Nado, 2014b). As such, I take it that use of the term 'intuition' picks out a diverse set of mental states or events, rather than a single one with any specific or unique characteristics.

It might be clear to the reader that AG fails to lend any support to the Irrelevance Claim. Again, the reason for this is that the AG at most lends support to ID. But, as previously discussed, this falls short of endorsing the Orthogonality Claim—i.e., the claim that intuitions play no epistemically significant role whatsoever in philosophy. But, could we reformulate AG to lend support to the Orthogonality Claim? For the remainder of this section, I argue we cannot.

1.5.2 The revised argument

The AG easily can be charitably reformulated so as to putatively motivate the Orthogonality Claim. Recall, the original formulation of AG builds on the idea that philosophers' use of the term 'intuition' is defective: it does not denote any particular mental state or event. Now, note that this proposal can be taken to motivate the strong claim that considerations about the epistemology of intuitions fail to pick out *any* epistemically significant dimension of our practices of philosophical inquiry. Otherwise put, this amounts to the claim that since 'intuition' has no plausible referent, then talk of intuition does not refer to anything

that could be used in an epistemically significant way in philosophy. Thus, we should conclude that considerations about intuitions would not be significant for debates about the epistemology of philosophy. More schematically:

The Argument from Gibberish revised (AGr)

P1 Philosophers' talk of 'intuitions' is defective: it fails to denote any particular mental state or event.

P2 If philosophers' talk of 'intuitions' is defective, then it fails to pick out *anything* that could play an epistemically significant role in philosophy.

C Therefore, intuitions do not play any epistemically significant role in philosophy.

I contend that the AGr fails as there are good reasons to reject its second premise. To illustrate, consider the curious fact that until not too long, the two different minerals jadeite and nephrite were wrongly categorised as one single mineral called 'jade'. Now, note that the term 'jade' as used to refer to jadeite and nephrite is, on Cappelen's terms, 'semantically defective'. After all, the term 'jade' fails to secure a *specific* semantic content as it does not denote any *specific*, unique mineral. But, it is clear that even if 'jade' is thus semantically defective, this does not show that claims of the sort 'this necklace is made of jade' fail to pick out some material of which the necklace is made of. Of course, this claim still leaves it undefined just what the necklace is composed of—i.e., it is unclear whether it is made of jadeite or nephrite. But, this does not mean that speaking of 'a necklace made of jade' fails to denote *some* material that makes up this necklace.

These considerations carry over for the case of 'intuition'. Plausibly, even if 'intuition' turns out to be semantically defective, this does not mean that talk of intuitions fails to pick out *something* that plays an epistemically significant role in philosophy. This is compatible with the claim that talk of 'intuition' makes it unclear just what mental state or event that is. But, this is not tantamount to saying that there is no such mental state or event. Thus, I find there are good reasons to resist the second premise of AGr, and that this argument thereby fails to motivate the Orthogonality Claim.

1.6 Conclusion

In this chapter, I argued that Cappelen and Deutsch fail to establish that experimental findings about intuitions are wholly irrelevant for an assessment of the epistemology of philosophy (I dubbed this the 'Irrelevance Claim'). I first explained why Cappelen and Deutsch are wrong to think that the Irrelevance Claim follows from their arguments to the effect that philosophers do not use intuitions in an evidential role. I then proposed that if the arguments Cappelen and Deutsch advance are to stand any chance of vindicating the Irrelevance Claim, they must successfully motivate the claim that intuitions play no epistemically significant role whatsoever in philosophy (I called this the 'Orthogonality Claim'). I then argued that even under the most charitable reading of their arguments, Cappelen and Deutsch fail to

establish the Orthogonality Claim. In conclusion, I submit that Cappelen and Deutsch cannot vindicate the Irrelevance Claim, and that their arguments do not refute the experimentalist challenge.

Chapter Two

Testing for the Phenomenal

For the following two chapters, I consider a prominent line of reply to the experimentalist challenge according to which empirical studies about intuitions are in need of considerable refinement before they can be put to use in debates about the methodology of philosophy. In this chapter, I focus on a formulation of this proposal that builds on the thesis known as phenomenalism—i.e., the idea that intuitions have a *characteristic* phenomenology. Phenomenalists contend that since empirical studies about intuitions do not account for this phenomenology, then it is unclear whether their findings actually reflect evidence of the workings of intuitions. Thus, they argue that we simply lack sufficient empirical evidence to conclude that intuitions are sensitive to epistemically irrelevant factors, and that the experimentalist challenge is thereby unmotivated. I dub this the 'phenomenalist rejoinder'.

Here is a detailed plan of the chapter: I first rehearse the central arguments for the phenomenalist rejoinder (sec. 2.1). I then briefly expound recent developments from research on human metacognition and argue that a particular set of *metacognitive states*—viz., a group of mental states with a *high* feeling of rightness—can be aptly characterised as intuitions (sec. 2.2). As we will see, this proposal provides very good reasons to think that the phenomenalist rejoinder fails. In the last section (sec. 2.3), I argue that these developments in research on human metacognition pinpoint sources of error in intuitions and can thus help philosophers make better informed assessments of their warrant for relying on intuitions in inquiry.

2.1 Phenomenalism

2.1.1 The phenomenology of intuitions

The linchpin of all recent phenomenalist accounts is the idea that intuitions and perceptual experiences are fundamentally similar kinds of mental states. A first line of support for this idea is that ordinary language seems to suggest as much: For example, we often speak as if we can come to *see* by intuition that there cannot be four-sided triangles, or that intuitions make it *seem* to us that killing a person to save five others is immoral. Phenomenalists take these metaphors as illustrative that *what it feels like* to have an intuition is similar to *what it feels like* to have a perceptual experience (Bengson, 2015b; Chudnoff, 2013; Koksvik,

2011). In more detail, phenomenalists suggest that both intuitions and perceptual experiences possess a (broadly defined) *presentational phenomenology*.

Examples help to clarify what this "presentational phenomenology" of perceptual experiences amounts to. For instance, consider the difference between a case where a subject S *sees* that a wall is red and a case where she is *told* that it is red. Phenomenalists propose that only S's perceptual experience will have a phenomenology that makes it seem that the wall is red and that presents S with the *redness* of the wall. In more detail, they contend that only S's perceptual experience will make it seem to her as if she is aware of a relevant property of the wall (viz., its phenomenal redness) that itself motivates her to endorse that the wall is red. Furthermore, they suggest that intuitions also possess a presentational phenomenology. Thus, a subject S's intuition that *p* will both (i) make it seem to S as if *p* is true, and (ii) *present* S with relevant properties of *p* that motivate her to endorse *p* (Bengson, 2015b; Chudnoff, 2013; Koksvik, 2011). Again, examples are helpful to clarify. Consider the following two propositions:

Proposition 1 Two non-concentric circles have at most two common points.

Proposition 2 If a quadrilateral is inscribed in a circle, the sum of the products of the two pairs of opposite sides is equal to the product of the diagonals

Chudnoff expects that many people will have an intuition that Proposition 1 is true after they consider this claim (Chudnoff, 2013, pp. 50-51). In more detail, he predicts that by imagining and manipulating the mental imagery of two circles, it will (i) seem to readers as if Proposition 1 is true, and (ii) that they will be presented with the properties of circles that seem to make it true that they intersect in at most two points—which motivates the reader to endorse Proposition 1. By contrast, he expects that most people will not have an intuition about Proposition 2 when they consider this claim. That is, he predicts that thinking about a quadrilateral inscribed in a circle is unlikely to elicit a mental state in readers which both makes it seem to them as if Proposition 2 is true, and that presents them with the relevant properties of this figure that motivate them to endorse Proposition 2.

Bengson also seeks to delineate the contours of the phenomenology of intuition by contrasting them with other mental states. He proposes that the following anecdote is useful for this end:

> **Ramanujan's Intuition**: The mathematical prodigy Ramanujan is on his way to visit his mentor, Professor Hardy, in London. He hails a cab and, as it stops, he notices that its number is 1729. This causes him to smile, for he immediately sees that this number has a very interesting property, namely the property of being the smallest number expressible as the sum of two positive cubes in two different ways.
>
> (Bengson, 2015b, p. 711)

For most people, entertaining thoughts about the number 1729 will not elicit the kind of mental state which Ramanujan is reported to have. More precisely, most of us do not immediately recognise 1729 as

decomposing into two pairs of positive cubes—at least not in the same way most people would recognise that 2 + 2 = 4. In this sense, Ramanujan's mental state is distinctive and unlike our own: for him, it *seemed* true that 1729 has such a particular mathematical property. Surely, most of us may perform calculations which lead us to believe or appreciate that 1729 is the smallest number expressed by two pairs of cubes; but, thinking this is not identical to being *presented* with its truth as in Ramanujan's case. Bengson proposes that this putative difference between our mental states and that of Ramanujan upon considering the number 1729 is illustrative of what is distinctive of the phenomenology of intuitions.

Koksvik also make recourse to this approach of phenomenal contrasts to describe the phenomenology of intuitions. The crucial distinction Koksvik points to is that between two instances: an initial moment of not being aware of an impression regarding logical entailment of a given proposition, and a later of having such an impression:

> I am neutral about the truth or falsity of many propositions. There are many I have never considered, and many about which I regard myself to have no evidence either way. Let one of these be the proposition that *p*. [...] Let p be if my shoes are by the door, then they are not not by the door. Suppose I have never considered this proposition. But now I do, and it comes to seem to me that if my shoes are by the door, then they are not not by the door.
>
> (Koksvik, 2011, p. 177)

By comparing and contrasting the phenomenology of intuitions with that of other mental states, phenomenalists seek to establish the following two claims:

Presentational thesis (PT) All intuitions possess a presentational phenomenology.
Irreducibility thesis (IT) Intuitions are irreducible to any other mental state.

Phenomenalists claim that these examinations lend support to PT insofar as they reveal that intuitions within a host of different domains (mathematics, logic, and philosophy) share a same presentational phenomenal character. Furthermore, they argue that these examinations suggest IT as they reveal that other mental states lack this same phenomenology. More precisely, phenomenalists argue that there are four features of the phenomenology of intuitions that discern them from any other mental state (Bengson, 2015b; Chudnoff, 2013; Koksvik, 2011). I detail these below.

First, intuitions are **conscious, non-perceptual propositional attitudes** that incline assent to what they represent. In other words, they are intentional states that do not involve any of the sensory-modalities of perception, and which represent a proposition as true—thus motivating endorsement of that proposition. In this sense, Koksvik suggests that it only seems true to one that 'if the shoes are by the door, then they are not not by the door' once one consciously entertains the relevant proposition. Accordingly, if one had not entertained that proposition, then one would not undergo such an experience making it seem that this was so. Similarly, Chudnoff proposes that it is the conscious consideration of a proposition, and

manipulation of the relevant mental imagery regarding properties of circles, which itself gives rise to an impression as if "Two non-concentric circles intersect in at most two points" is true. Lastly, the anecdote about Ramanujan describes his experience as following his attending to the number 1729 on the taxi; it does not simply strike him that this is so, independent of any conscious awareness of this number.

Second, they are **spontaneous** in that we do not have to weigh considerations for or against a given proposition before having an intuition about it. For instance, we do not deliberate on whether circles do, or do not intersect in at most two points before having an intuition that this is the case. Rather, it just simply strikes us that this is so once we manipulate the mental imagery of two circles. For instance, an intuition about the number of intersections between two non-concentric circles does not require one to explicitly entertain additional propositions about the mathematical properties of circles so as to assent on its truth. On a similar note, one does not conclude that 'if one's shoes are by the door, then they are not not by the door', by weighing considerations in favour or against the truth of this claim—it merely seems to one that this is so. As an additional point, consider the Gettier intuition 'Smith does not know who will get the job'. It is implausible to hold that this intuition ensues as a result of conscious consideration of theoretical frameworks regarding the necessary and sufficient criteria for knowledge attributions.

Third, intuitions are **autonomous**, insofar as they are not responsive to contrary evidence. That is, it is possible for a subject S to have an intuition as if *p*, even if she reflectively endorses, believes, or judges that *not-p*. To illustrate this point, phenomenalists note that many people report having a persistent intuition that the naïve comprehension axiom[1] is true, even though they *know* it to be false.

And lastly, intuitions are **seemingly truthful**, in that they purport to make subjects aware of facts. In this sense, Chudnoff suggests that the reader's intuition that **Proposition 1** is true will seem to make them aware of *why* it is true. In more detail, he proposes that this intuition will enable thoughts about the properties of how circles intersect, which make it true that they have at most two common points (Chudnoff, 2013, pp. 50-51). On a similar note, Bengson suggests that intuitions make subjects aware of features of their contents that seem to rationalise assent, "in the (psychological) sense that they tend to make formation of corresponding beliefs seem rational or fitting from the first-person perspective" (Bengson, 2015b, p. 723). Thus, he proposes that intuitions make it seem 'from the inside' as if their contents are 'worthy of belief'. Likewise, Koksvik suggests that, part of what it is like to have an intuition is that it "purports to represent an objective fact" (Koksvik, 2011, p. 168).

Now, phenomenalists do not intend for these four features to give an exhaustive account of the phenomenology of intuitions. Nevertheless, they take them as *sufficient* to demarcate intuitions from any other kind of mental state (Bengson, 2015b; Chudnoff, 2013; Koksvik, 2011). For instance, they suggest that intuitions are markedly distinct from doxastic states (e.g., beliefs and judgements), insofar as the latter are *not* autonomous—since we believe, judge, or are inclined to believe or judge in accordance with the evidence available to us. They also argue that intuitions are unlike mere guesses or hunches, which

[1] The naïve comprehension axiom states that for every arbitrary condition *x*, there will be a set containing all, and only the things meeting that condition *x*.

they contend are *not* seemingly truthful—as they *at most* incline one to believe or judge, without making their contents seem worthy of belief. And so on and so forth for all other kinds of mental states, none of which, they propose, possess all the above four features.

Critically, for the purposes of this chapter I will grant both PT and IT. That is, I take it that a presentational phenomenology is a marker of intuitions, and that an analysis in terms of the above four features suffices to distinguish intuitions from any other mental states. Accordingly, I henceforth use the unqualified term "intuition" to refer to mental states with a presentational phenomenology. In this way, I aim to give the most charitable interpretation of the phenomenalist views and of their reply the experimentalist challenge, which I now turn to.

2.1.2 A rejoinder to the experimentalist challenge

The proposal that intuitions are defined by way of their phenomenology motivates a forceful critique of the experimentalist challenge. Central to this reply is a dispute over a common assumption operative in empirical studies motivating this challenge: Namely, the proposal that people's answers on questionnaire-style surveys amount to reports of intuitions. However, if the phenomenalist thesis is warranted, it suggests that this assumption is questionable. After all, mere reports of a person's answers are not instructive as to whether they had an experience with a presentational phenomenology when responding to that task. Bengson (2013) provides a very clear formulation of this proposal. Commenting on the studies motivating such methodological concerns, he suggests that:

> [T]hese attacks neglect a considerable gap between the answers elicited by the relevant empirical studies and the intuitions about which naysayers naysay. It cannot innocently be assumed that subjects' answers expressed how things struck them—what intuitions they had, if any. The point is simple, but not insignificant. For, I will argue, it implies that we are at the present time unwarranted in drawing any negative conclusions about intuitions from the relevant empirical studies.
>
> (ibid., p. 496)

To buttress this proposal, phenomenalists examine prominent studies in experimental philosophy with a view to teasing out how their findings might conflate intuitions with other mental states—for example, hypotheses, guesses, emotional reactions, inferences, or other conclusions (ibid., p. 497; Chudnoff, 2013, pp. 107-113). For instance, Bengson contends that the studies showing *wording effects* on the Trolley problem might be running together intuitions with mere emotional reactions. Specifically, he contends that because such studies highlight the negative outcome of the scenario (i.e., that one person will be killed), they trigger negative emotions, which leads people to refrain from opting to change the course of the trolley (Bengson, 2013, pp. 517-518). Building on such examinations, phenomenalists suggest that there are decisive reasons to think that empirical studies about intuitions are simply inadequate to motivate the experimentalist challenge.

In the next section I argue that there are good reasons to think that this phenomenalist rejoinder fails. As we will see, recent inquiry into human metacognition has studied a class of psychological states that share crucial phenomenological commonalities with intuitions. I then discuss evidence that these psychological states display strikingly similar patterns of variation as those uncovered by studies in experimental philosophy. Lastly, I explain why these findings undercut the phenomenalist rejoinder to the experimentalist challenge. To set up this argument, I start out in the next section by summarizing key features of recent research on human metacognition.

2.2 Metacognition and Phenomenalism

Imagine someone asks you "Who wrote *A Clockwork Orange*?". You stop for a moment and try to recall; you are sure that you know the answer, but the name just escapes you now. Or suppose you are given a piece of chapter with a couple of math problems. You try to solve the first one, but you feel as if you have got it wrong. You move on to the next, which looks much easier. You then become pretty confident that you got it right. Common to all these events is the operation of human capacities of metacognition: The set of capacities that allow for representation and monitoring of one's own cognition (Proust, 2013).

Much of the recent research on metacognition has focused on the study of so-called *metacognitive experiences*—a number of phenomenal states that are *about* one's own cognition (Arango-Muñoz and Michaelian, 2014). For current purposes, I focus on the metacognitive experience known as the *feeling of rightness* (henceforth FOR; FORs for the plural "feelings of rightness"). This refers to a gradable impression that the answer arrived at in response to a cognitive task is correct. Low degrees of FORs yield a weak impression of accuracy and are correlated with higher rates of change in later responses. By contrast, high degrees of FORs give a strong impression that the answer produced is accurate, which make people less likely to change their responses later on (Thompson, Turner, and Pennycook, 2011; Thompson, Turner, and Pennycook, 2013; Thompson and Johnson, 2014).

Notably, there are at least very superficial similarities between mental states with a *high* FOR and intuitions: Namely, both have a phenomenology which makes their contents seem *true* to a subject. In what follows, I flesh out the similarities between intuition and mental states with FOR to defend the following claim:

Int.FOR Some mental states with a *high* FOR are intuitions.[2]

In support of Int.FOR, I first argue that a particular group of mental states with a high FOR display precisely the features that phenomenalists take as sufficient to demarcate intuitions from other mental states (detailed in the previous section). I then detail why Int.FOR suggests that phenomenalist rejoinder to the experimentalist challenge fails.

[2] This thesis is not to be read as a bi-conditional. Thus, I leave it open whether every intuition is a mental state with a high FOR.

2.2.1 Phenomenalism and high FOR

Investigations into the workings of the FOR make use of a number of reasoning tasks to elicit and examine this metacognitive experience (Thompson, Turner, and Pennycook, 2011; Thompson, Turner, and Pennycook, 2013). These experimental studies take on a very similar approach to the one that phenomenalists use to argue for their views: Namely, they invite readers to entertain particular examples, and to then immediately attend to the mental states that arise. To illustrate the type of task used in research on the FOR, consider the following pair of examples:

Ex. 1 If a car runs out of gas, it will stall. The car has not stalled, therefore it has not run out of gas.

Ex. 2 None of the nurses are magicians. Some of the winemakers are nurses. Therefore, some of the magicians are not winemakers.

Alternatively, consider the following two examples:

Ex. 3 Suppose you are told that a particular pub follows the rule that if a person drinks alcohol, then they must be at least 18 years old. Of the four people at the pub, the first is 16 years old, the second is 18 years old, the third is drinking beer, and the fourth is drinking water. To check whether the pub is actually enforcing their own rule, you need only ask what the 16-year old is drinking, and whether the person drinking beer is more than 18 years old.

Ex. 4 Most puppies have a price. Some friendly animals have a price. Some puppies are friendly animals.

Research on the FOR suggests that examples like Ex. 1 and Ex. 3 are likely to elicit mental states with a high FOR in the reader. Thus, readers who consider Ex. 1 are expected to immediately accept that the car would indeed not have run out of gas, and for it to seem to them as if it is *true* that this conclusion follows from the premise. Similarly, reading Ex. 3 is predicted to make it seem true that one need only ask the 16-year old what they are drinking, and how old is the person drinking beer, to check whether the pub is enforcing the rule. Thus, evaluating Ex. 1 and Ex. 3 will elicit experiences that are categorically distinct from a mental state with a low FOR such as a mere *guess* or *hunch*, which can at most incline accepting or rejecting a statement, but which does not make it seem true. Conversely, Ex. 2 and Ex. 4 are unlikely to elicit a mental state with a high FOR: Most readers who entertain Ex. 2 and Ex. 4 will not immediately arrive at an answer that accompanies a feeling as if this response is accurate.

These brief illustrations underscore superficial similarities between intuitions (as these are defined by phenomenalists) and mental states with a high FOR: Namely, both make their contents seem true. However, further careful reflection on these examples of high FORs brings to light some other very significant and important commonalities with intuitions. Most notably, these mental states with a high FOR have a phenomenology that is very similar in kind to that which phenomenalists ascribe to intuitions. For instance, I expect that considering Ex. 1 will both make it seem to readers as if this claim is true and that they will be presented with that which makes it true. That is, this mental state will make it seem

to readers as if they are *aware of properties of* Ex. 1 that motivate endorsing this claim—for example, that Ex. 1 is just a case of modus tollens, which makes it true that the car would not have run out of gas. Likewise, I expect that entertaining Ex. 3 will make this claim seem true to the reader and will present them with properties about this claim that motivate endorsing Ex. 3. In this sense, it will seem clear to readers *why* one need only ask what the 16-year old is drinking and the age of the person drinking beer to check whether the bar is following its own rule.

Furthermore, there is additional support for the parallel between intuitions and the above examples of mental states with a high FOR in that the latter are also aptly characterized in terms of the four features that phenomenalists take as sufficient to demarcate intuitions from all other mental states. First, the mental states with high FOR above are conscious non-perceptual propositional attitudes that incline assent to what they represent. That is, the mental states that ensue from considering both Ex. 1 and Ex. 3 are intentional states that represent a proposition as true, thus motivating endorsement of that proposition, but which do not involve the sensory modalities of perception. Second, these examples of mental states with a high FOR are also *spontaneous*: They simply occur without the need to weigh considerations for or against a certain proposition. Thus, I expect that readers who consider Ex. 1 will not have to deliberate about any one position or other about whether a car has run out of gas before having a mental state with a high FOR that Ex. 1 is correct. Rather, it will simply strike readers as true that the car has not run out of gas. Third, these mental states with a high FOR are *autonomous* in the sense that they do not respond to contrary evidence. Indeed, a large body of empirical research suggests that people are prone to have mental states with a high FOR about claims which they do not judge or believe to be true (Ackerman and Zalmanov, 2012; Thompson, Turner, and Pennycook, 2013; Thompson, Evans, and Campbell, 2013).

And lastly, these mental states with a high FOR are *seemingly truthful* in that they purport to make subjects aware of facts. In line with Chudnoff's explanation of what this amounts to, I propose that a mental state with a high FOR about Ex. 1 will seem to make one aware of *why* this proposition is true. For instance, I expect that when the reader considers Ex. 1, it will seem to them as if they are made aware of the precise logical properties instantiated by Ex. 1 which make it true—for example, readers will recognize that Ex. 1 is a case of modus tollens. Furthermore, in line with Bengson's proposal, I submit that identifying Ex. 1 as a case of modus tollens will make it seem rational from the reader's perspective to form a corresponding belief that Ex. 1 is true. On a similar note, I predict that when the reader considers Ex. 3, it will seem to them as if they are made aware of *why* this claim true. In this sense, I expect that the reader will be made aware of the logical relations that make it true that one need only ask two of the people at the bar to check whether the pub is following its own rule, and that this will make a belief in Ex. 3 seem rationally fitting from the reader's perspective.

This additional analysis suggests that the mental states with a high FOR that ensue from considering Ex. 1 and Ex. 3 share a set of important features with intuitions. Specifically, like intuitions, they are also aptly described as being *conscious non-perceptual* propositional attitudes that incline assent to their contents, and which are *spontaneous, autonomous, and seemingly truthful* in character. Moreover, to the extent that intuitions are characterized by way of these features, this analysis suggests that the mental states with a

high FOR elicited above can be aptly described as intuitions—that is, they lend support to the thesis I have called Int.FOR. So, in what follows, I consider the mental states with a high FOR that arise from considering Ex. 1 and Ex. 3 to be intuitions.

For the remainder of this section, I assume Int.FOR as a working hypothesis and proceed to tease out its implications. In particular, I argue that this thesis shows there to be substantial empirical support against the phenomenalist rejoinder. To set up this argument, I will now detail evidence that mental states with a high FOR of a similar kind to those exemplified above display patterns of variation that are strikingly similar to those uncovered by studies in experimental philosophy.

2.2.2 Experimental findings

A central ambition of studies into the workings of the FOR has been to identify the determinants of this metacognitive experience. Researchers have thus conducted studies with a view to examining whether people's reports of FORs vary with respect to changes in certain defined parameters. Many of these studies report that mental states with a high FOR display patterns of variation that are surprisingly similar to those uncovered in studies from experimental philosophy. Specifically, they provide evidence of variation in people's reports of high FORs depending on how questions and tasks were framed (Thompson, Turner, and Pennycook, 2011), the order in which experimental tasks were presented to them (Markovits, Thompson, and Brisson, 2015, exp. 2), as well as demographic factors, such as whether people were of high cognitive capacity or low cognitive capacity (Thompson and Johnson, 2014). I will now first briefly describe one such set of findings—namely, those showing that mental states with a high FOR very similar to Ex. 1 and Ex. 3 (discussed above) are subject to a variety of *framing effects*. I then build on these and other findings of worrying patterns of variation in mental states with a high FOR to argue that the phenomenalist rejoinder to the experimentalist challenge fails.

Thompson and colleagues (2013) report evidence of framing effects on mental states with a high FOR on the famous Wason Selection Task (henceforth, WST). We have already seen an example of this task. That is because Ex. 3 is one version of the WST. On more traditional formulations of the WST, people are presented with pictures of four cards and informed that each card has a letter printed on one side, and a number on the other. Two of the cards display the side printed with a number, and the other two display the side with the letter (e.g., the cards could be labeled as "A", "S", "3", and "7"). People are then asked to read a conditional statement expressing a rule about the relation between the number and the letter of a specific card (e.g., "If a card has the letter 'A' on one side, it has the number '3' on the other"). Then, they are asked to identify which of the cards shown need to be turned so as to verify whether the set of cards being shown are compatible with this rule—that is, if they do not violate this rule.

Most people give the wrong response on this traditional—so-called "implicit negation"—version of the WST. For instance, when asked which cards would verify the rule "If a card has the letter 'A' on one side, it has the number '3' on the other", people either choose to flip the cards showing the "A" and the "3", or just the card displaying the "A". (The correct answer is in fact to select the cards labeled "A" and "7".) However, people are much less prone to giving this wrong answer on an explicit negation version of the

WST. On this version of the task, all the cards are labeled with either the letter or number mentioned in the rule or their negation—such that the rule "If a card has a letter 'A' on one side, it will have a '3' on the other" would be followed by cards labeled "A", "not-A", "3", or "not-3" (Evans, Clibbens, and Rood, 1996; Stahl, Klauer, and Erdfelder, 2008). And similarly, people also tend to give the right answer to the WST, when this task is formulated in terms of everyday social situations, such as in Ex. 3.

Thompson and colleagues (2013) conducted an experiment aiming to further examine this shift in people's responses to the WST. In this study, they asked people to complete either an implicit or explicit negation version of the WST within a novel two-response paradigm used to study the FOR. They asked people to first issue a quick answer to the WST and evaluate their degree of confidence about this response; in a second experimental phase, they asked people to reconsider their first answer in free-time and to evaluate their confidence in this second response. In line with previous findings, Thompson et al. (ibid.) also provide evidence that people in the implicit negation condition often gave the wrong answer to the WST. More importantly though, they found that these people tended to also attribute a high FOR about this answer and to endorse this answer when later asked to reconsider it. Conversely, people in the explicit negation condition were more likely to give correct responses to this task and to attribute high FORs to their answers—also often endorsing that answer when asked to reconsider. In effect, these findings thus reveal a framing effect on mental states with a high FOR about the WST. Specifically, they reveal that minor changes in the superficial features of the WST (viz., framing it either in terms of implicit or explicit negations) can determine the answers to which people report high FORs.

In another study, Thompson and colleagues (2011) provide additional evidence of framing effects in mental states with a high FOR. In this set of experiments, they asked people to evaluate what logically followed from a simple conditional statement of the form "If A then B",again using the two-response paradigm described above. We have already seen one case of this kind of task above in Ex. 1. As in Ex. 1, all the conditional relations used in this set of experiments made reference to familiar objects and situations from everyday life; however, for half of the questions used, the conclusions were coded as being "unbelievable" in that they were likely to contradict people's background knowledge about the objects and situations there mentioned. For instance, the following is a case of such an unbelievable conclusion: "If the TV is plugged in, then it works; the TV is plugged in, therefore, it works". Most people know that TVs can fail to work even if plugged in (they may be faulty, or power may be out). Knowledge of this fact makes the above conclusion strike most people as implausible ("unbelievable"), even if that conclusion does follow logically from the premises. The remaining conclusions in the experiment were framed so as to be made believable in that they were compatible with background beliefs about the items mentioned in the task—for example, "If a car runs out of gas, then it will stall; if the car has not stalled, therefore it did not run out of gas".

Thompson et al. (ibid.) found that participants were more likely to evaluate that believable conclusions followed from the premises and that they frequently attributed high FORs about this response, even when the inference was invalid. However, people often judged unbelievable conclusions to *not* follow from the premises and to also report high FORs about these responses, even when the inference was valid (for

similar findings: Thompson, Turner, and Pennycook, 2013). In effect, this pattern of evaluations pinpoints a framing effect in mental states with a high FOR. More precisely, they show that framing conclusions as either believable or unbelievable can have a significant impact on whether people have a high FOR about whether it follows or not from a given set of premises.

Besides these two sets of findings, additional studies have uncovered similarly worrying framing effects on mental states with high FORs in evaluations of conditional statements—similar to that detailed in Ex. 1 (Ackerman and Thompson, 2017). Furthermore, inquiry into the FOR has also found other worrying patterns of variation in mental states with a high FOR to similar tasks among people with high or low cognitive capacity (Thompson and Johnson, 2014), and the order in which cases are presented to people (Markovits, Thompson, and Brisson, 2015; Markovits et al., 2017).

I contend that, when taken together, these studies raise significant doubts about the phenomenalist rejoinder. Recall, this amounts to the suggestion that the experimentalist challenge is ultimately unmotivated, insofar as we lack sufficient evidence to conclude that intuitions—rather than a host of other mental states—vary with respect to the parameters identified in studies in experimental philosophy. But, I have argued that the mental states with high FOR that arise from considering Ex. 1 and Ex. 3 can be aptly defined as intuitions. Furthermore, I have argued that the empirical literature shows that the mental states with high FOR like Ex. 1 and Ex. 3 vary with respect to similar kinds of factors to those used to motivate the experimentalist challenge.[3] In light of this, I submit that there is mounting evidence that the phenomenalist rejoinder fails.

2.2.3 Objections

I now turn to a brief exposition and rejection of what I take to be the two most promising replies available to phenomenalists against the arguments in this section. The first is to claim that findings about mental states with a high FOR discussed above are inadequate to assess the phenomenalist rejoinder. This reply builds on the observation that the examples used to elicit these mental states with a high FOR involve only simple conditional statements or arguments. However, these examples are not the kinds of cases that are typically used in the philosophical literature to elicit people's intuitions—for example, Gettier scenarios, or the Trolley problem. And since the phenomenalist rejoinder refers to the use of intuitions in *philosophy*, then there are reasons to doubt that the findings discussed above actually concern the intuitions that are at stake in the experimentalist challenge. As such, phenomenalists might claim that there is little reason to think that those findings show that *philosophical* intuitions are prone to vary in problematic ways.

Although seemingly plausible, there are significant difficulties for this line of reply. The central issue is that it relies on tracing a significant distinction between philosophical and non-philosophical intuitions. However, phenomenalists maintain that all intuitions—regardless of their content (e.g., mathematical, logical, philosophical)—are categorized by way of their shared presentational phenomenology, which demarcates them from any other mental state. On this view, there are no intrinsic and fundamental differences among intuitions at the level of their subject matter that would warrant claims to the effect that

[3] For additional empirical support, see also Danek and Wiley (2017).

philosophical intuitions are more reliable than any other. As such, I propose that the only option available to phenomenalists to salvage this first line of reply would be to claim that the mental states with a high FOR I have examined are not in fact intuitions. I take it that the most promising way to flesh out this proposal is to argue that, since the examples used to elicit those mental states with high FOR involve evaluations of the validity of conditional statements, then they are actually *inferences*—not intuitions. This then raises doubts about the attempt to make use of findings about FORs to put pressure on the phenomenalist rejoinder.

However, this alternative way of articulating the first reply fails by the phenomenalists' own lights. That is because phenomenalists themselves have advanced arguments to the effect that considering conditional statements can give rise to an intuition about their validity—such that it will seem to one that it is valid or not, and make it seem that one is *presented* with that which makes it so (Chudnoff, 2013, pp. 149-150; Koksvik, 2011, p. 177). For example, Chudnoff suggests that considering the following simple argument:

(1) Every even number is divisible by two.
(2) The number of pigs in the pen is even.
(3) So, the number of pigs in the pen is divisible by two.

can give rise to "an intuition experience that represents that (1) and (2) support (3)", which will seem to *present* one with why this is the case (Chudnoff, 2013, p. 150). Likewise, Koksvik (2011, p. 177) contends that considering the claim, "If my shoes are by the door, then they are not not by the door" will give rise to an intuition that *presents* this inference as valid. Similarly, I propose that the kinds of mental states with a high FOR examined in this chapter are naturally described as intuitions of just this sort: That is, they are intuitions about the logical validity of an inference. After all, as argued for above, these mental states with a high FOR do possess the characteristic (presentational) phenomenal character of intuitions, and are thus aptly defined as intuitions. Given these claims, we then return to the first point: If such intuitions about the logical validity of conditional statements are supposed to be markedly distinct from philosophical intuitions—for example, such that only the former are prone to the worrying patterns of variation found above—then phenomenalists would owe us an explanation for why this should be the case. Even so, I should note that it is quite difficult to see just why intuitions about the logical validity of inferences are in any way philosophically irrelevant. For example, consider the kinds of informal arguments that philosophers routinely advance to the effect that a given theory implies a rather unpalatable conclusion, which is taken to suggest that the theory is thereby false. In such cases, it is very reasonable to expect that philosophers often rely on intuitions about what logically follows from the central claims of that theory to evaluate whether that theory does indeed imply that problematic consequence. In this sense, I take it that intuitions about the logical validity of inferences are relevant for debates surrounding the experimentalist challenge (see sec. 2.3 for additional discussion of this point). For these reasons, it is difficult to argue that the empirical findings about mental states with high FOR discussed above are inadequate to evaluate the phenomenalist rejoinder.

A second promising line of reply to the arguments in this section is an objection analogous to the phenomenalist rejoinder, yet aimed at studies on the FOR. The main thrust of this type of reply is the idea that findings from these experimental studies are likely to conflate mental states with high FORs with a variety of other mental states, such as people's post-hoc rationalizations, beliefs, guesses, and hunches. Building on this proposal, phenomenalists might object to the suggestion that findings from inquiry into the FOR are apt to motivate methodological concerns about the use of intuitions in philosophy.

Although seemingly plausible, a main difficulty for this objection is that empirical studies on the FOR have made considerable efforts to avoid conflating mental states with a high FOR with other such confounding phenomena. In particular, they make use of a novel two-response experimental paradigm that has been found effective for this end (Thompson et al., 2011). On the first experimental stage of this framework, people are asked to quickly give an answer to reasoning problems and to then immediately rate their FOR about this initial answer (marking this on a scale from one to seven). Adoption of this first quick-response paradigm is motivated by empirical findings which show that asking people to give their first immediate judgements about a task interferes with their ability to engage in deliberate reflection about the experimental task at hand (Neys, 2006), to prompt less neural activation from areas associated with belief inhibition (Tsujii and Watanabe, 2010), and to lead people to report those answers that immediately strike them as plausible—rather than what they infer or suppose is the correct response to that task (Evans and Curtis-Holmes, 2005). This first stage of the experimental design thus helps researchers home in on mental states with high FORs by minimizing the possibility that people's responses reflect their background beliefs, inferences, and suppositions about that task.

The second-stage of the two-response paradigm used in inquiry into FOR is similarly well motivated. On this second stage, people are asked to reconsider their initial response to the experimental task—now in free-time. A robust finding from studies on FOR is that people's reports of a high FOR about their answer to an experimental task is a very good indicator that they will endorse that answer when asked to reconsider it.[4] Owing to this, mental states with high FORs are regarded as distinct from a simple hasty or unreflective response—of the sort one would quickly reject upon closer scrutiny. This indicates that the second stage slow-response paradigm is also suitable to inquire into the FOR. In particular, to the extent that this allows researchers to track changes in people's first and second answers, and to correlate them with reports of high FORs, this helps to distinguish between mental states with high FORs from other confounding phenomena—such as mere hasty or unreflective responses.

In addition, studies on the FOR have also actively sought to rule out the possibility that their findings might reflect people's mere guesses or hunches about the experimental task (Thompson, Turner, and Pennycook, 2011; Thompson, Evans, and Campbell, 2013). As previously mentioned, both of these are defined as mental states with a low FOR, insofar as they at most incline subjects to assent to a claim without making it seem as if that claim is true. In order to rule out interference of guesses and hunches in their findings, researchers instructed participants in their experiments to report FORs on a Likert scale ranging from one to seven, in which the lower ends of this scale were clearly labeled as a mental state

[4] For a review of the evidence, see Ackerman and Thompson (2017)

with a low FOR—for example, "Just guessing". In effect this means that people's reports of high degrees of FOR in these experiments were made in explicit contrast to self-reports of mere guesses or hunches.

Taken together, these considerations raise significant difficulties for the claim that studies about the FOR might conflate mental states with a high FOR with a variety of other phenomena. As detailed above, the two-stage experimental paradigm used in inquiry into FOR has been shown effective at minimizing the possibility that experimental findings might reflect evidence of participants' post-hoc rationalizations, beliefs, judgements, inferences, and mere guesses or hunches. As such, the concerns that phenomenalists have raised about studies in experimental philosophy do not naturally carry over to studies on the FOR.

2.2.4 Summing up

I have been arguing that some mental states with a high FOR can be aptly defined as intuitions (Int.FOR), and that these mental states are prone to patterns of variation that are strikingly similar to those uncovered in studies in experimental philosophy (e.g., framing effects, the order in which cases are presented, and demographic variables). I then proposed that these findings indicate that intuitions do indeed vary in the ways that experimental philosophers have suggested, and that this refutes the phenomenalist rejoinder to the experimentalist challenge. Lastly, I discussed and rejected the most promising replies available to phenomenalists against these arguments.

Now, one important upshot of these arguments is that they provide further support to the experimentalist challenge. After all, they both undercut a prominent objection to these experimentally motivated methodological concerns, and give additional evidence that intuitions vary with respect to truth-irrelevant factors. However, as mentioned in the outset of this thesis, it is still a matter of great dispute just what these findings imply about the use of intuitions in philosophy. Proponents of *radical* versions of the experimentalist challenge propose that these findings show that philosophers should refrain from using them in inquiry. Proponents of the *moderate* experimentalist challenge disagree: they suggest that appeals to intuitions in philosophy are sometimes warranted, but that these should be accompanied by a better understanding of when we can trust intuitions, and under what circumstances they are prone to leading us astray.

For the remainder of this chapter I argue that Int.FOR lends support to a moderate experimentalist challenge. To develop this proposal, I first briefly describe two of the most prominent formulations of *radical* versions of the experimentalist challenge—namely, those proposing that intuitions are *unreliable* or *hopeless*. I then argue that both these formulations fail.[5] As we will see, the arguments I advance against these views will also reveal how findings from inquiry into the FOR pinpoint sources of errors in intuitions, which inform philosophers how to make better use of them in inquiry.

[5] It is noteworthy that there are other ways to respond to radical formulations of the experimentalist challenge. For instance, Williamson (Williamson, 2007) has forcefully argued that such formulations of the experimentalist challenge quickly overgeneralise to scepticism about any judgement whatsoever (for a reply: Nado, 2015). Alternatively, some have suggested that such formulations are self-defeating, as they themselves rely on intuitions to articulate their central claims (for an insightful discussion: Srinivasan, 2015).

2.3 Advancing the Debate

2.3.1 Radical experimentalist challenges

A very common reading of findings from experimental philosophy is that they speak to the reliability of our intuitions. In this sense, evidence that particular intuitions vary with respect to truth-irrelevant factors are taken to show that *those* intuitions are unreliable. More controversially, many experimental philosophers regard these worrying findings as illustrative of the epistemic deficiencies potentially afflicting other philosophically-relevant intuitions as well. Thus, they suggest that the charge of unreliability extends much wider than just these local demonstrations—e.g., to most intuitions about ethics (Sinnott-Armstrong, 2011), epistemology (Swain, Alexander, and Weinberg, 2008), philosophy of mind (Feltz and Cokely, 2012), and, more generally, to all philosophical intuitions (Machery, 2017, ch. 3). Given the sensible assumption that we should refrain from using unreliable sources of evidence in philosophical inquiry, then the above considerations suggest that we should not make use of intuitions in philosophical inquiry.

There are, however, good reasons to resist this strong conclusion. Specifically, a growing body of empirical work suggests that a number of philosophically relevant intuitions stem from cognitive processes that are in fact generally reliable. For instance, Jennifer Nagel (2012a) has argued that intuitive knowledge attributions that are central to epistemology arise from the exercise of ordinary psychological capacities for "mind-reading"—that is, a set of operations humans routinely rely on to understand and predict the mental states of others. Given the considerable amount of evidence that capacities for mind-reading are very reliable—as shown by the accuracy of people's predictions of others' mental states—Nagel (ibid.) suggests that it is rather sensible to expect that intuitive knowledge attributions will be reliable too (see also Boyd and Nagel, 2014). A set of studies by Eugen Fischer and colleagues motivate a similar conclusion. In a number of experiments, they report evidence that philosophical intuitions about perceptual experiences are underwritten by processes that are key to our competence as speaker-hearers of a language—such as metaphor interpretation (Fischer, 2014) and use of stereotype-driven inferences in verb comprehension (Fischer and Engelhardt, 2016; Fischer and Engelhardt, forthcoming). Moreover, Fischer et al. contend that since many philosophical intuitions are elicited by verbal descriptions of hypothetical cases (e.g., the Trolley-problem, or Gettier cases), then it is likely that those intuitions are underwritten by similar cognitive processes of linguistic comprehension. Again, as there is substantial evidence that such processes are generally reliable (as shown by how we successfully use them in everyday communication) then there are good reasons to think that they generate reliable intuitions as well.

The central upshot of the above considerations is that they significantly weaken the inductive step that takes us from local demonstrations of unreliability to the claim that philosophically-relevant intuitions are in general unreliable. Thus, they undercut methodological concerns about the use of intuitions in philosophy that build on this proposal.[6] However, it is noteworthy that even if we grant that intuitions are in fact generally reliable, this is still insufficient to show that philosophers are in any way warranted in

[6] Machery (2017) has recently attempted to respond to this objection from evidence of reliability. Although I find there is some merit to his reply, I find that the arguments I develop in subsequent chapters (3 & 4) provide good reason to deny that they suffice to motivate a radical version of the experimentalist challenge.

appealing to them in inquiry. For instance, Jonathan Weinberg (2007) has argued that what is at issue in methodological concerns about the use of intuitions is whether they are "hopeless". Weinberg suggests that a source of evidence provides warrant only if it is, in his technical sense, "hopeful", by which he means that we are able to both identify and correct for that source's errors when, and if, these arise (ibid., p. 327). Furthermore, he suggests that intuitions fare quite badly in this regard insofar as we simply lack an adequate understanding of when intuitions are prone to lead us astray, and what we can do to mitigate such errors. Moreover, although Weinberg does not endorse a radical experimentalist challenge, these considerations can easily lend support to this view: For, given Weinberg's claims that only "hopeful" evidential sources can provide warrant, and that intuitions are ultimately "hopeless", then this clearly implies that intuitions are not apt to provide any warrant for philosophical positions and claims.

For the remainder of this chapter, I argue that the claim according to which intuitions are ultimately "hopeless" fails to account for the many insights that research on the FOR provides into the workings of intuitions. As we will see in the next section, findings from this research helpfully single out particular factors that are prone to problematically influence people's intuitions and can thus help philosophers identify and mitigate for such errors when they arise. Thus, the findings from inquiry into the FOR both undercut the above *radical* experimentalist challenge and help to articulate a moderate version of this view.

2.3.2 A moderate experimentalist challenge

I have previously discussed evidence that the perceived believability of an inference's conclusion can influence people's intuitions about logical validity. In more detail, these findings show that people often judge inferences with "believable" conclusions to be valid, even when they are in fact invalid. Conversely, people are much more likely to judge inferences with "unbelievable" conclusions as invalid when they are in fact valid. One constructive way of reading these findings is that they helpfully pinpoint vitiating circumstances that can lead intuitions astray and which philosophers would do well to be wary of. Thus, they suggest a fairly simple set of practical recommendations that can potentially help improve intuition-based methodologies of philosophy: Take measures to minimize, or rule out, the possibility that people's judgements of validity might reflect *merely* that which they find to be believable.

One way to lend weight to this proposal is to show how it proves apt to explain and help mitigate a worrying pattern of variation uncovered in a recent study by Nichols and Knobe (2007). In this study, they report evidence of a robust framing effect in people's intuitions about moral responsibility and free will. They first presented people with the following description:

> Imagine a universe (Universe A) in which everything that happens is completely caused by whatever happened before it. This is true from the very beginning of the universe, so what happened in the beginning of the universe caused what happened next, and so on right up until the present. For example, one day John decided to have French Fries at lunch. Like everything else, this decision was completely caused by what happened before it. So, if

> everything in this universe was exactly the same up until John made his decision, then it had to happen that John would decide to have French Fries
>
> (ibid., p. 669).

People in the abstract condition were then asked "In Universe A, is a person fully responsible for their actions?". Conversely, people in the concrete condition were asked to first read the description of Bill—an individual from Universe A who murders his wife and children and then runs away with his secretary—and were then asked "Is Bill fully morally responsible for his actions?". Nichols and Knobe found that people answering the abstract question often judged that a person in Universe A would not be responsible for their actions, although people who read the case of Bill judged him morally responsible for murdering his family.

Evidence that the believability of conclusions might sway people's intuitions about logical validity can aptly explain this difference in evaluations. To begin spelling this out, it is first helpful to get clear on the structure of the above experimental task. In particular, note that judgements to both the abstract and concrete questions are essentially evaluations of the validity of two variations of a logically identical conditional statement. Specifically, the abstract question asks readers to evaluate whether it follows that "If Universe A obtains, then any person in this universe is morally responsible for their actions". Similarly, people in the concrete condition are in effect asked to evaluate whether it follows that "If Universe A obtains, then an individual in this universe called Bill (who murdered his family) is morally responsible for his actions". We can describe the difference in evaluations in terms of the validity of inferences: Whereas people in the abstract condition judged the conditional statement as invalid, people in the concrete condition judged it to be valid.

Now, consider how the consequent of the conditional statement that mentions Bill is aptly described as being believable insofar as it focuses on the act of murder. After all, most of us strongly believe that murderers should be held responsible for their actions. Conversely, the conclusion of the more 'abstract' conditional can be described as unbelievable, as it focuses on no action in specific—and most of us do not believe that people should be held responsible for actions when they really had no other choice. Thus, the evidence that people tend to judge inferences with believable conclusions as valid and unbelievable conclusions as invalid seems apt to explain the above difference in evaluations. In this sense, people's judgement that it follows that "If Universe A obtains, then an individual in this universe called Bill (who murdered his family) is morally responsible for his actions" can be attributed to the believability of this conclusion. Conversely, evaluations that it does not follow that "If Universe A obtains, then any person in this universe is morally responsible for their actions" is due to the fact that this conclusion is unbelievable. This then highlights that the proposed methodological recommendations are actually instructive. In particular, they illustrate a case in which philosophers would do well to attend to the potential deleterious effects of believable/unbelievable conclusions in intuitions about validity.

We can tease out some further helpful methodological suggestions from another set of findings from inquiry into the FOR. As previously discussed, Thompson and colleagues (2013) report evidence of

framing effects on intuitions in the WST. In more detail, they found that people who complete the implicit negation version of this task—in which only two cards have lexical content matching those of the rule under examination—often choose to examine precisely the cards bearing superficial similarities to the rule. Conversely, people who complete the explicit negation version of this task—in which *all* cards have lexical content matching the rule—often choose a distinct set of cards. Thompson et al. (2013) suggest that this finding illustrates how the similarity in the lexical content of questions and reasoning prompts can make particular elements appear more relevant to the task at hand, and thus induce people to focus on them when issuing answers. Again, we can read these findings as providing instructive methodological guidance for philosophers in that they point out that people's intuitions about philosophical relevant scenarios might be swayed by lexical similarities between questions and the descriptions of cases. Thus, they suggest the following practical recommendation: Try to minimize, or rule out, the possibility that people's intuitions might reflect *merely* their examinations of what is made more salient by the wording of the question.

To bolster this proposal, I show how it aptly explains a set of recent findings from a study by Alexander and colleagues (Alexander et al., 2017). In this study, they report evidence that people's intuitions about many typical cases used in philosophical discussions about peer-disagreement display significant framing effects. For instance, in one experiment they presented people with the following variation of the very well-known "Restaurant Case":

> Suppose you and your friend go out to dinner. When it is time to pay the check, you agree to split the check evenly and to give a 20% tip. You do the math in your head and become highly confident that your shares are $43 each. Meanwhile, your friend does that math in her head and becomes highly confident that your shares are $45 each. You and your friend have a long history of eating out together and dividing the check in your heads, and know that you've been equally successful at making these kinds of calculations: usually you agree; but when you disagree, you know that your friend is right as often as you are. Moreover, you are both feeling sharp tonight and thought that the calculation was pretty straightforward before learning that you disagreed about the shares.
>
> (ibid., p. 2540)

One group of people were then asked the question "Should you give your friend's belief equal weight and think that it is no more likely that you are right than that your friend is right, or should you continue to prefer your own belief?". Another group was asked instead, "How confident should you be that your belief is correct now that you know that your friend disagrees with you?". Alexander and colleagues (ibid.) report that people in the first group often judged that they should reduce confidence in their initial answer in this case, whereas people in the second group responded that they would maintain highly confident in it.

One plausible explanation for this finding is that the words used to formulate these questions highlight distinct elements of the Restaurant Case, thus prompting very different examinations of the scenario. To

begin developing this proposal, note how the first question invites people to consider giving "equal weight" to a friend's response, and asks whether their "friend is right". In light of the evidence detailed above, it is plausible that this formulation highlights precisely the aspects of the Restaurant Case that are worded in a similar way: Thus, it brings attention to the description of the friend as "equally successful" and that "your friend is right as often as you are" in these kinds of mental calculations. In this sense, this particular framing of the question underscores the parts of the case that emphasize just how reliable the friend actually is; as such, it is quite unsurprising that people answering this first question often judged they should reduce confidence in their own belief in light of the disagreement with their friend.

Now contrast this with the formulation of the second question in which people were asked to rate how "confident" they should be in their answer given that the friend "disagrees with you". Again, given the evidence of framing effects due to similarity in lexical content, it is plausible that this framing is likely to highlight those aspects of the case that are similarly worded. As such, they bring attention to the description of how the reader is "highly confident" in their calculation, which is described as being pretty straightforward besides the fact that the friend "disagreed about the shares". As such, this particular formulation makes salient just how strongly the reader believes their answer to a fairly straightforward calculation is correct, and that only that friend disagrees. In this light, it is also unsurprising that people answering this question chose to maintain their high degree of confidence in their answer; after all, most of us, when we are very confident about what we believe, will not concede so easily when challenged—as testified by the endless disputes between people with different opinions.

This explanation lends support to the methodological suggestion advanced above. On this interpretation of the findings from Alexander et al. (ibid.), the framing effects on the Restaurant case are caused by commonalities in the wording of this scenario and the follow-up questions people responded to. This shows that, at least with respect to this case, philosophers should not ask follow-up questions which use lexical content matching those from the description of the relevant scenario, so as to avoid swaying people's intuitions.

Lastly, it is worth mentioning how the body of empirical work on the FOR is suggestive with respect to a recently very influential proposal due to Jennifer Cole Wright (2010). She contends that attending to the perceived strength of one's own intuitions is a reliable method for tracking the evidentiary value of philosophical intuitions. This proposal is motivated by empirical evidence that people in experimental studies report being less confident about intuitions that are unstable—that is, those that display the worrying patterns of variation uncovered by experimental philosophers (for similar results: Zamzow and Nichols, 2009). This suggests that philosophers should restrict their appeals to those intuitions about which people tend to report greater levels of confidence, since these are less prone to such deleterious effects.

Research into mental states with FORs shows the above methodological suggestions to be misguided. As discussed in previous sections, these studies have uncovered evidence that mental states with a high FOR—that is, responses about which people reported a high degree of confidence—are also prone to the worrying patterns of variation found in studies in experimental philosophy. Thus, people's sense of

confidence about their own intuitions is not a trustworthy method to weed out those intuitions that are subject to deleterious effects. Instead, I propose in line with the above considerations that philosophers should pay closer attention to those specific factors that have been found to sway people's intuitions (e.g., believability of conclusions and similarities in lexical content).

In sum, I contend that we can tease out useful methodological suggestions from inquiry into the FOR that help philosophers make better use of intuitions in philosophy. Relying on recent findings in experimental philosophy, I illustrate two cases in which these suggestions prove effective at mitigating framing effects in philosophical intuitions. In light of this, I take there to be good grounds on which to deny the radical versions of the experimentalist challenge that build on the claim that intuitions are ultimately "hopeless". Furthermore, I propose that these considerations thus also help to articulate a moderate version of the experimentalist challenge insofar as they inform philosophers of a set of vitiating circumstances that helps them assess their warrant for relying on intuitions in inquiry.

2.4 Conclusion

In this chapter, I have argued that developments from research on human metacognition can help to make headway in a set of very thorny methodological disputes about the role of intuitions in philosophy. Relying on findings from this body of empirical work, I argued that a class of metacognitive states can be aptly characterised as intuitions, and that empirical findings about these undermine a prominent objection to the experimentalist challenge. I then showed how inquiry into human metacognition provides us with a better understanding of the vitiating circumstances that can potentially lead intuitions astray, and demonstrated how they offer useful methodological suggestions that help philosophers make better use of intuitions in philosophy. These arguments thus offer valuable resources that can greatly contribute to the improvement of philosophical methodology and are suggestive with respect to future inquiry into the nature and epistemology of intuitions.

Chapter Three

Philosophical Expertise Under the Microscope

In this chapter, we[1] consider a second prominent formulation of the claim that extant empirical studies about intuitions fail to motivate the experimentalist challenge. According to this line of reply, the relevant empirical studies are problematic insofar as they take non-philosophers to compose their participant pools. However, some Defenders claim that philosophers are *experts*, such that their intuitions are likely to be shielded from the deleterious effects found to afflict the intuitions of non-philosophers. This has become known as the 'expertise defence'. This chapter explores the viability of the expertise defence, and argues that there is empirical evidence to suggest that it can be partially vindicated.

Here is a detailed plan of the chapter. To set up our argument, we first briefly review the two most prominent approaches to formulating the expertise defence and show how they operate with a problematic assumption: namely, that the expertise defence either stands or falls *wholesale* (sec. 3.1). We argue that the many ways in which intuitions are put to use in philosophical inquiry engender distinct *types* of philosophical expertise, and that wholesale defences or rejections of the expertise defence are thereby unwarranted—instead, we argue that local, piecemeal investigations of philosophical expertise are better suited to assess the expertise defence (sec. 3.2). In more detail, we suggest that the expertise defence must be sensitive to the cognitive underpinnings of intuitions appealed to within philosophy (sec. 3.2.1), and to different implementations of thought-experiments (sec. 3.2.2), and argue that neglecting these distinctions significantly inhibits our thinking about the expertise defence (sec. 3.2.3). Finally, in the spirit of taking our own advice, we demonstrate how recent developments in *cognitive epistemology* enable a piecemeal assessment of philosophical intuitions and, crucially, provides empirical support for at least one instance of a successful expertise defence (sec. 3.3).

[1] In this chapter, I adopt a collective voice ('we') as the ideas in this chapter were developed in collaboration with Lewis D. Ross (who contributed to the elaboration of roughly 20% of the article on which this chapter is based).

3.1 Current Approaches to the Expertise Defence

The expertise defence is primarily motivated by the observation that most experimental critiques of intuitions rely on data gathered from studying non-philosophers. Defenders point out that we might naturally think philosophers are just better equipped than non-philosophers to engage in philosophical inquiry. If so, then the force of the experimentalist challenge rests on an illicit inference: that failures affecting the judgements of non-philosophers generalise to the judgements of *expert* philosophers. Although this response has a ring of plausibility, it requires considerable precisification. We will organise our discussion around two independent claims that any version of the expertise defence needs to establish:

1. Philosophers have *philosophical expertise*—adequate training in philosophy leads to improved performance in some task(s).
2. Philosophical expertise *defends* philosophers from distorting factors that Critics claim undermines the practice of appealing to intuitions in philosophy.

So, what could the nature of the expertise be that philosophical training brings about, and how might it help philosophers avoid deleterious biases? In this section, we review and critique two natural responses to this question. Firstly, we outline and reject the 'expert intuitions' defence—the view that locates philosophical expertise in the superior intuitions of philosophers. Secondly, we outline and critique the more plausible 'expert practices' defence, which locates philosophical expertise in the superior methods employed by philosophers.

3.1.1 Expert Intuitions

Call the suggestion that philosophical expertise yields superior intuitions the *expert intuition defence*. In line with the criteria outlined above, this amounts to the following proposal:

Expert Intuition Defence: Philosophical training leads philosophers to have intuitions that are substantially less susceptible to the distorting effects influencing the intuitions of non-philosophers.

There are several ways to flesh out this basic suggestion. For instance, Kauppinen, (2007) proposes that philosophers bring a distinctive kind of *reflective* intuition to bear on philosophical cases. Ludwig, (2007) argues that philosophers are apt to provide intuitions that stem from *conceptual competence* alone, rather than responses influenced by other factors (such as pragmatic implicatures). Both versions of Expert Intuitions contrast the intuitions at play in philosophical discourse with the quick, spontaneous responses provided by non-philosophers to experimental surveys. Their hope is that the more refined intuitions of philosophers respect fine-grained distinctions elided by the folk, whilst stripping out the influence of various irrelevant factors and routine inattentiveness to cases.

Although it might seem promising to deny the symmetry between the intuitions of philosophers and quick judgements participants provide in experimental surveys, Expert Intuitions Defence is a difficult position to maintain. Firstly, there is substantial empirical evidence that many of our judgements in a variety

of domains are susceptible to errors due to the effect of cognitive biases. If the influence of biases are a pervasive part of human psychology, susceptibility to bias in philosophical discourse—like in other domains—is a fair default assumption. Furthermore, Critics have already gathered a small but growing body of recent evidence suggesting that trained philosophers display no particular immunity to cognitive biases (for reviews: Machery, 2015; Mizrahi, 2015)). Without countervailing empirical evidence, it is hard to see how we could be justified in holding the sort of exceptionalism Defenders wish to attribute to professional philosophers.[2] Biases are prevalent in all sorts of judgements and specific measures are required to ameliorate their influence; believing against the empirical evidence that philosophers' intuitions are somehow immune to these biases seems to be an article of faith rather than a well-supported position.

3.1.2 Expert Practices

The issues facing Expert Intuitions have led some to conclude that preoccupation with intuitions *simpliciter* is misguided. Jennifer Nado succinctly captures this worry:

> [The] focus on intuition is somewhat odd, given that it is in no way essential to the expertise defence; philosophical methodology obviously consists in much more than the gathering of intuitive judgements, and there are therefore many potential loci for philosophical expertise.
>
> (Nado, 2014a, p. 1029)

Nado goes on to suggest that philosophers "display expertise in assessing, critiquing, and (when appropriate) rejecting or explaining away" the data gathered by the method of cases (ibid., p. 1041).[3] Williamson, (2011) also shows dissatisfaction with the narrow focus on intuitions; in this sense, he characterises philosophical expertise as the improved performance in the *use* of thought experiments in philosophical inquiry—by which he means the evaluation, assessment, and deployment of them in theorising (ibid., pp. 224-225).[4] And there are interesting yet brief comments along similar lines to be found in Rini, (2015, p. 3). We take these comments to suggest philosophical expertise is to be found in what philosophers *do* with intuitions. As an attempt to provide a rough locus for our discussion that remains neutral on various metaphilosophical positions, we will characterise the variety of philosophical practices that might be promising for the expertise defence as follows:

Philosophical Practices: Philosophers elicit, invoke, assess, and synthesise intuitions about philosophical cases.[5]

[2] See also Weinberg et al., (2012)

[3] It is noteworthy that Nado endorses an account of philosophical expertise that does not support the expertise defence.

[4] Neither Nado nor Williamson phrase their accounts in terms of 'intuitions', arguing that intuition-talk is theoretically inert and often leads to confusion. However, our use of 'intuitions' here is roughly equivalent to 'judgements elicited from philosophical cases', which both Nado and Williamson refer to in their accounts.

[5] We borrow this description of how philosophers use intuitions in philosophy from Fischer and Collins, (2015, p. 11)

Attempts to locate philosophical expertise in what philosophers *do* with intuitions is a significant improvement on previous intuition-focused proposals of the expertise defence: we agree with Nado that there is more to philosophy than simple reliance on intuition, and that this broader nexus of practices must be taken into account when examining the nature of philosophical expertise.

However, we argue that extant formulations of this approach remain problematic. Specifically, we suggest that these attempts to refine the locus of philosophical expertise are problematically *monolithic*. By this, we mean that current approaches imply that the philosophical expertise relevant to **Philosophical Practices** consists of: (i) *a single package of* (ii) *general philosophical skills* that are (iii) *shared by all/most philosophers*, and are (iv) *apt to be successfully deployed across a range of philosophical contexts*.[6] This monolithic view has important ramifications for how we think about the expertise defence which, recall, is roughly composed of two claims: 1) Philosophers have *philosophical expertise*—adequate training in philosophy leads to improved performance in some task(s); and 2) Philosophical expertise *defends* philosophers from distorting factors that critics claim undermines the practice of appealing to intuitions in philosophy. The monolithic account of philosophical expertise currently implied in the literature leaves us with two basic possibilities:

Expert-Practices Defence: Philosophers display improved performance in the practices of eliciting, invoking, assessing, and synthesizing intuitions for purposes of philosophical inquiry. Moreover, *this suffices* to shield philosophers from mistakes and problems afflicting non-philosophers when appealing to intuitions in philosophical inquiry.

No Expert-Practices Defence: Philosophers display improved performance in the practices of eliciting, invoking, assessing, and synthesizing intuitions for purposes of philosophical inquiry. However, *this does not suffice* to shield philosophers from mistakes and problems afflicting non-philosophers when appealing to intuitions in philosophical inquiry.

These two options capture the tendency—prevalent in the literature—to either defend or refute the expertise defence *simpliciter*. That is, both sides in the debates about the expertise defence advance accounts which say that, whatever the relevant form of philosophical expertise consists in, it either *univocally does* or *does not* support the expertise defence (Williamson, 2007; Kauppinen, 2007; Ludwig, 2007; Weinberg et al., 2010; Williamson, 2011; Tobia, Buckwalter, and Stich, 2012; Nado, 2014a; Machery, 2015; Rini, 2015).

[6] There has been some recognition in the literature of the need to distinguish between types of 'philosophical expertise' when investigating the expertise defence. For instance, Weinberg et al., (2010, p. 335) recognise that there is an important distinction to be drawn between improved performance in "close analysis of texts, or the critical assessment of arguments" and improved performance "at conducting thought experiments". In a similar vein Nado, (2014a, p. 1041) distinguishes expertise in the "construction of thought experiments" and reflecting on philosophical theories, viz. "teasing out their consequences, determining their compatibility with other positions, and so forth." After drawing such distinctions, Nado and Weinberg are concerned with asking whether these types of expertise might plausibly shield philosophers from bias; on this front, they find little grounds for optimism. However, when philosophical expertise is put in this way, the expert practices defence begins to look rather unpromising. After all, even if philosophers are better than the average person at analysing texts, or coming up with intriguing thought-experiments, this doesn't seem apt to protect them against the multiplicity of biases identified by Critics.

We propose that any such extreme pessimism or optimism is unwarranted. Specifically, we will argue that this way of construing the expertise defence rests on an inadequate conception of philosophical expertise that neglects important distinctions among tasks encompassed by **Philosophical Practices**. We advance an alternative option, not identified in the literature, that aims to be more faithful to the complex reality of such practices. The main thrust of our proposal is the idea that the relation between philosophical expertise and the expertise defence is considerably more intricate than the binary choice between **No Expert-Practices Defence** and **Expert-Practices Defence**. More precisely, we argue for the following:

Expert-Practices Defences: *Some* philosophers display improved performance in *certain sets of practices* in eliciting, invoking, assessing, and synthesizing intuitions for purposes of philosophical inquiry. Moreover, this suffices to shield *those* philosophers from *specific* mistakes and problems afflicting non-philosophers when appealing to intuitions in philosophical inquiry.

In the following section, we motivate **Expert-Practices Defences** by arguing that the diversity of practices described under **Philosophical Practices** engender distinct *types* of philosophical expertise. In this sense, we suggest that philosophical expertise in such practices consists of: (i) *a diverse package* of (ii) *specific philosophical skills* that are (iii) *shared by some philosophers*, and are (iv) *apt to be successfully deployed only in a narrow range of philosophical contexts*. To articulate this view, we detail two methodologically important distinctions that are overlooked in current discussions of philosophical practices: first, between the types of cognitive processes underlying philosophical intuitions, and second, among the different ways in which thought-experiments can be put to use in philosophical inquiry. We then demonstrate why eliding these distinctions fails to enable an informative evaluation of the expertise defence.

3.2 Finessing the Expertise Defence

3.2.1 Cognitive Processes

In recent years, there has been a growing consensus that intuitions do not form a homogeneous class of mental states.[7] This claim is rather plausible: there do seem to be significant differences between, for instance, intuitive judgements about whether a subject knows some proposition, and whether pushing people in front of trains for the greater good is morally permissible. Given the superficial dissimilarities between these intuitions, it is quite natural to wonder whether the practices subsumed under **Philosophical Practices** will fare equally well with respect to each of these intuitions. We will argue that they do not, and that discussion of philosophical practices would benefit from taking into account the differences between distinct intuitions.

Recent developments in experimental philosophy provide empirical evidence suggesting that philosophical inquiry should distinguish between different types of philosophical intuitions. The studies in question employ techniques and findings from the empirical sciences with the aim of identifying the particular

[7] Nado, (2014b) provides a very clear discussion of these points.

cognitive processes underlying our philosophical intuitions.[8] For current purposes, we need only detail developments from two such lines of inquiry.

The first concerns recent proposals that seek to trace back a certain class of epistemic intuitions to the operation of a set of specific cognitive processes. The epistemic intuitions in question concern intuitive knowledge-attributions: judgements about what other people may or may not know under certain circumstances—e.g., whether people can know the time from a stopped clock, even if it is only by sheer luck that the clock indicates the correct time at that very moment. In a set of recent papers, Jennifer Nagel (Nagel, 2012b; Nagel, 2012a; Boyd and Nagel, 2014) offers a psychological explanation of these intuitive knowledge-attributions. Making use of a large body of findings from cognitive psychology, Nagel proposes that these intuitions stem from the exercise of the ordinary psychological capacities for 'mind-reading'—i.e., a set of automatic cognitive processes that humans routinely rely on in order to understand and predict the mental states of others.

The second strand of research we wish to call attention to concerns recent attempts to identify the cognitive underpinnings of intuitions about whether inflicting harm can ever be morally permissible, if doing so would bring about some greater good. Most empirical inquiry into the nature of these intuitive judgements seeks to pinpoint the specific processes driving people's responses to the so-called 'Trolley Problem'—i.e., cases where people are asked if it would be morally permissible to in some way alter the course of a run-away trolley in order to save many people, at the cost of sacrificing the life of one person (for a discussion: Thomson, 1985). A number of empirical investigations into the nature of these intuitions provide robust neuroscientific evidence suggesting that the intuition according to which it would *not* be morally permissible to alter the course of the trolley—thus sacrificing one to save the many—are driven by a *negative emotional response* to this case (for a review: Greene, 2014, pp. 700–706). Thus, these studies suggest that our verdicts in response to the 'Trolley Problem' are shaped by the affective processes we bring to bear when considering it.

These empirical findings provide evidence that there are in fact significant dissimilarities between the many types of intuitions that philosophers appeal to in philosophical inquiry. Specifically, they illustrate how different intuitions can be driven by distinct cognitive processes. Now, one may ask why any of this is relevant for an evaluation of the expertise defence. Here's why: these discoveries reveal how different intuitions are susceptible to distinctive kinds of *biases*. To explain, let's first turn to empirical evidence of the biases afflicting intuitive knowledge-attributions.

One hotly debated topic in contemporary epistemology concerns whether the standards for knowledge are sensitive to the error possibilities that are made salient in a given context (DeRose, 1992; Cohen, 1999). A central issue driving these debates is the observation that people are usually more willing to attribute knowledge to others when no possibilities of error are made salient; however, once such possibilities are raised, most people feel the temptation to retract such attributions. This amounts to a startling pattern of judgements, given that the orthodoxy has it that merely mentioning the *possibility* of error should

[8] Indeed, as Knobe (2016) suggests, most contemporary work in experimental philosophy is practised as cognitive science.

not shift the standards for knowledge. Building on psychological explanations of intuitive knowledge attributions in terms of 'mind-reading' capacities, Nagel (2012b) makes a move for the orthodox view. Specifically, she argues that these patterns of intuitive judgement are due to a specific *psychological bias* known as *epistemic egocentrism*, or the *curse of knowledge*. Simplifying, this bias amounts to a tendency to wrongly predict the mental states of other subjects who stand in epistemic positions that are more naive than one's own. In these situations, most people erroneously represent others' perspectives as if they shared their own privileged epistemic standing—and, moreover, they tend to penalise others (e.g., retract knowledge attributions) for not responding to the evidence as they think appropriate from their privileged epistemic standing. Subsequent empirical investigations have borne out the predictions of this psychological explanation, providing additional evidence that changes in intuitive knowledge-attributions in the light of possibilities of error can indeed be explained in terms of the egocentrism bias (Alexander, Gonnerman, and Waterman, 2015).

Let us now turn to a different set of debates: namely, discussions about the Doctrine of Double Effect (DDE)—i.e., the principle which states that it is sometimes permissible to bring about harm as a *side-effect* of aiming to bring about some good, even though it would be impermissible to bring about the same harm as a *means* of bringing about the same good. One of the central lines of support for the DDE is that it neatly accounts for typical responses to certain complex morally charged scenarios. For instance, this principle nicely captures one common response-pattern to two variants of the Trolley-Problem:[9]

Lever: A runaway trolley is headed toward five innocent people who are on the track and who will be killed unless something is done. Patrick may pull a lever, which will redirect the trolley onto a second track, saving the five people. However, on this second track is an innocent bystander (Susy), who will be killed if the trolley is turned onto this track.

Push: A runaway trolley is headed toward five innocent people who are on the track and who will be killed unless something is done. Patrick can push an innocent bystander (Susy) in front of the trolley. The runaway trolley would be stopped by hitting (and killing) Susy, but would thereby save the five people on the track.

According to the DDE, it would (i) be permissible to stop a train from killing five people by pulling a lever to *divert* the train onto a second track, leading to the death of one person, but (ii) it would not permissible to *push* one person onto the track—thereby killing them, but stopping the train—in order to prevent the train from killing five people. Indeed, this is precisely the pattern of responses most people have when they evaluate these cases. But here's the rub: studies show that people can be primed to give different judgements to **Lever** and **Push**, simply by modifying the order in which these cases are presented to them (Schwitzgebel and Cushman, 2012; Schwitzgebel and Cushman, 2015). For instance, if one is presented with the **Push** case first then you are more likely to judge the subsequent **Lever** case to be impermissible. One prominent explanation for why people are susceptible to such order effects is the operation of an

[9] These formulations are modified from Liao et al., (2012).

affective bias: one's judgements are being prompted by an unreflective consideration of one's emotional state. By using one's immediate emotional reaction as a benchmark, we become disposed to changes in intuitive response depending on whether we are presented with an emotionally 'potent' variation of the thought-experiment (such as pushing a person to their death) in the first instance or later on (for a discussion: Greene, 2014).

What is the upshot of all this? It shows that particular biases and errors afflicting certain intuitions are not likely to arise for others. And as a result, different skills will be required to successfully appeal to an expertise defence within some lines of philosophical inquiry than in others. For instance, an expertise defence for philosophers working specifically on the Trolley-problem will require them to deploy practices that can shield them from certain affective biases[10] (in order to either prevent the biases, or to ameliorate their effects on their intuitions) whilst another philosopher working on the relationship between knowledge and stakes will have to be cautious regarding biases such as epistemic egocentrism. These are skills that they require in virtue of their lines of inquiry engaging different cognitive processes.[11] Therefore, asking whether expert philosophical practices suffice for an expertise defence cannot be carried out in an intuition-general way; the sorts of biases one is susceptible to in a given line of inquiry, and the skills and practices required to offset these biases, will depend on the cognitive processes producing one's intuitions.

To take stock, we have briefly documented recent developments in experimental philosophy indicating that our intuitions are underwritten by very distinct kinds of cognitive processes. Furthermore, we explained how these psychological explanations of our intuitions reveal them to be vulnerable to a variety of distinct biases. We then argued that these distinctions are methodologically significant, insofar as they suggest that distinct skills are required to offset biases afflicting different intuitions. These findings provide an initial attack on the monolithic interpretation of philosophical expertise—which holds that expertise in the tasks under **Philosophical Practices** amount to one set of general skills.

In what follows, we underscore yet another important methodological distinction that the monolithic interpretation fails to account for—viz., between different ways of implementing thought-experiments in philosophical inquiry—and outline its relevance for assessing the expertise defence. We then conclude this section by discussing an alternative approach to delineating philosophical expertise that can accommodate for these distinctions, and which allows for a more informative evaluation of the expertise defence.

3.2.2 Different uses of thought-experiments

Philosophers use thought-experiments in a variety of different ways. However, these distinct uses have not received equal scrutiny in the methodological literature. Here are three rather different ways that

[10] Of course, this isn't to say the ethical judgements should be free from affective influences *tout court*. We only make the uncontroversial point that, since order effects are clearly undesirable, ethicists must ensure that affective influences do not leave them vulnerable to such a deleterious bias.

[11] Again, this distinction cannot neatly be drawn by looking at *areas* of philosophy. For instance, it is noteworthy that some studies have found that considerations with moral valence can influence knowledge attributions (Beebe and Buckwalter, 2010). This implies that it is at least possible that affective processes may shift epistemic attributions, too. These initial results further support our suggestion that our focus should be on the specific cognitive processes our inquiries engage, not on the philosophical domain we take ourselves to be working in.

philosophers appeal to thought-experiments in philosophical inquiry (these distinctions are certainly not exhaustive):

Standard Uses: Firstly, and most prominently, philosophers use thought-experiments to *elicit judgements which are used to defend or refute claims* about concepts, folk theories, and/or the phenomena such concepts and folk theories are about. These uses of thought-experiments, what we are calling Standard uses, are widespread and familiar within the philosophical literature. The canonical examples are Gettier-cases. For most people, considering the details of Gettier-like scenarios cues a particular judgement—namely that subjects lack knowledge in certain 'lucky' conditions—and this judgement, in turn, is frequently used to defend a negative answer to the philosophical question 'is knowledge reducible to justified true belief?' Other famous examples of Standard uses of thought-experiments include Kripke's (1980) use of the Gödel case against descriptivist theories, Parfit's (1984) use of fission cases to argue against certain theories of personal identity, and Jackson's (1982) use of the Mary case to argue against Physicalism. In all these cases, the application of the thought-experiment seeks to elicit specific judgements, and then uses these judgements to *answer* some philosophical question.

Aporetic Uses: Secondly, some thought-experiments are used in order to elicit *conflicting judgements* about the same scenario. We will call these Aporetic uses of thought experiments, to indicate that they are meant to give rise to a kind of *puzzlement* that helps to refine inquiry into their subject-matter.[12] A simple example is the case of the ship of Theseus: "is a ship that has had all of its components gradually replaced on the course of a voyage the same ship on arrival as on departure?" Most people feel the pull of both the positive and negative answers: in one sense, it is indeed the same ship; but, in another, it seems that by substituting all of its parts we now have an entirely different ship (for discussions: Smart, 1972; Scaltsas, 1980; Lowe, 1983). This thought-experiment thus gives rise to a sense of puzzlement about the notion of identity, which can then be put to use in refining inquiry. For instance, if it turns out that the intuitive pull of each answer corresponds to a different concept of identity, then philosophers will have to hone their inquiry accordingly. Newcomb's problem is another instance of an Aporetic use of a thought-experiment. This is a puzzle that elicits two plausible yet inconsistent judgements about rational attitudes (Nozick, 1969).[13] Reflection on this problem has led to animated (and ongoing) discussions about the source of the puzzlement; it has promoted philosophers to draw novel distinctions and formulate questions that attempt make inquiry into rational attitudes more precise. (For instance, Newcomb's paradox shows that there are cases where the decision-theoretic principle of dominance gives advice that conflicts with the principle of expected utility-maximization; this is a surprising result as it dissociates our judgements about rationality from the predictions of standard decision-theory.) In sum, Aporetic uses of thought-experiments significantly improve philosophical inquiry, to the extent that they tease out important distinctions that help us to better understand central philosophical notions.

Critical Uses: A third way to use thought-experiments is to invoke particular intuitions in order to

[12] One philosopher who has discussed the paradoxical nature of certain thought-experiments is Sorensen (1992).

[13] As a recent poll suggests, people tend to diverge almost evenly on both judgements: from 31,854 votes, 53.5% chose one judgement and 46.5% chose the other (`https://www.theguardian.com/science/alexs-adventures-in-numberland/2016/nov/30/newcombs-problem-which-side-won-the-guardians-philosophy-poll`).

scrutinise which aspects of the specified case elicit them. Insights into which features of a case are responsible for eliciting certain intuitions can then be used in assessing the evidential value of those intuitions. Critical uses of thought-experiments are not predominant in the philosophical literature, yet they are arguably amongst the most discerning. One such example is Gendler's (2010) close scrutiny of a range of thought-experiments about personal identity. From this detailed analysis, she concludes that some of the most prominent thought-experiments used in the literature are just too far divorced from the situations in which we ordinarily apply the concept of personal identity. Her central contention is that our intuitions about such cases are thereby unlikely to track features essential to the concept of personal identity, making them useless for philosophical theorising. Another example of a Critical use is Jackson's (2003) *later* appeal to the case of Mary the Neuroscientist, where he attempts to diagnose *why* this case tends to elicit the intuition that Mary does indeed learn something new when she sees the red rose for the first time. Jackson invokes this intuition as a foil for his argument, which aims to show that this judgement turns on a mistaken conception of the nature of perceptual experience—viz., one which erroneously characterises experience in terms of *what it's like*. On these grounds, he proposes that we ought to reject this intuition, and that the case of Mary the Neuroscientist is inadequate for philosophical purposes.[14]

These three different ways to conduct philosophical inquiry have not—as far as we know—been separated out in the literature on the expertise defence. Critics have overwhelmingly focused on Standard uses, marshalling experimental evidence in order to argue that the practice of appealing to intuitions in this way is massively susceptible to truth-irrelevant factors. However, separating out different uses of thought-experiments reveals there to be a number of *different* practices that philosophers use in dealing with intuitions about cases. In other words, there is no *single* practice of eliciting, invoking, assessing and synthesising intuitions. The distinctions between Standard, Aporetic, and Critical illustrates at least some differences in how these practices can be put to use in philosophical inquiry.

The relevance of these distinctions is that these practices may have different epistemic profiles with respect to the influence of deleterious biases. To see this, consider how different uses of thought-experiment are used to inquire into the same philosophical issue—the nature, and perhaps the varieties of, personal identity. For example, Parfit's fission cases can be used as an argument for a particular view within debates about personal identity, and Gendler's use of various thought-experiments was used to reject the relevance of certain intuitions for theorising about personal identity. All of these turn out to be cases of eliciting, invoking, assessing and synthesising judgements about cases, but each for rather different purposes. Even if it were the case that Critics could adduce empirical evidence to demonstrate that Standard usages within this inquiry are susceptible to the influence of truth-irrelevant factors (e.g. order effects, affective biases, etc.), this would not be evidence that the Critical use of this thought-experiments are similarly susceptible. After all, it is plausible that Critical implementations of these thought-experiments could pick up on exactly these biases.

[14] Another prominent example of a Critical use of thought-experiments is Dennett's (1991, pp. 437-440) critical examination of Searle's Chinese Room scenario. Additionally, the recent literature in experimental philosophy provides examples of empirically-informed Critical uses of thought-experiments: e.g., Nagel, (2011) and Fischer, (2014).

Although this in itself is an interesting hypothesis for the dialectic surrounding the expertise defence, more will be required to push back against the experimental Critic. Specifically, it will need to be shown that certain uses of thought-experiments can in fact *shield* philosophers from the effect of biases and truth-irrelevant factors. In the final section, we will discuss recent empirical evidence that suggests that a Critical use of a particular thought-experiment, applied to one particular line of inquiry, dos indeed succeed in uncovering and shielding philosophers from the effect of bias and truth-irrelevant influences that have been found to affect the Standard use of that same thought-experiment in that line of inquiry.

3.2.3 Discussion

In this section, we have called attention to two important methodological distinctions that have been neglected in evaluations of the expertise defence. In particular, we suggest that the dominant monolithic interpretation of philosophical expertise does not adequately accommodate for distinctions between the various uses of thought-experiments in philosophical inquiry, nor the particular cognitive processes underlying philosophical intuitions.[15]

In line with the arguments in this section, we suggest that **Expert-Practices Defences** is a better way to think about the relation of support between philosophical expertise and the expertise defence. Recall, this is proposal that *some* philosophers display improved performance in a *particular set* of practices in eliciting, invoking, assessing, and synthesising intuitions, and that this suffices to shield *those* philosophers from *particular* deleterious effects on philosophical theorising. This proposal better accommodates the fact that philosophical expertise with respect to **Philosophical Practices** can stand in a variable relation of support to the expertise defence. Recall, the reason for this is that the many distinctions amongst the **Philosophical Practices** engender distinct types of philosophical expertise, some of which may shield philosophers from biases, whilst others do not.

The central upshot of **Expert-Practices Defences** is that sweeping conclusions of the Critics' and Defenders' arguments are unwarranted: they are illicitly motivated by the simplistic way that philosophical expertise has been related to the expertise defence. Acknowledging the variable relationship between philosophical expertise and the expertise defence renders invalid any inference taking us from a small sample of evidence for the expertise defence in one set of cases to a wholesale rejection or acceptance of it in all cases. Given these considerations, it is possible to remain open-minded about the expertise defence in advance of close study into the specific area and practice to which it is being applied.

However, an important objection to such agnosticism is in the offing: even if Critics cannot automatically draw large-scale negative metaphilosophical conclusions from their empirical studies, they might claim that these studies—along with the lack of countervailing evidence—justifies a strong presumption against the viability of any version (local or global) of the expertise defence. In order to give some succour to Defenders, we will now outline one example where robust empirical evidence *supports* a successful

[15] We should note that we do not argue that philosophical expertise should be carved only at the joints of these distinctions; rather, we merely illustrate some ways that current accounts fail to accommodate the possibility that different types of philosophical expertise.

instance of the expertise defence. We think that this approach of engaging with Critics on empirical grounds is more edifying than arguing over where the burden of proof lies in metaphilosophical debates; ideally Defenders and Critics will settle their metaphilosophical differences by gathering and evaluating targeted bodies of empirical evidence, not by attempting to claim temporary victories through dialectical fiat.

3.3 A (very) narrow account of the Expertise Defence

3.3.1 Cognitive Epistemology

One of the most promising recent developments in experimental philosophy is the 'cognitive epistemology' movement.[16] Proponents of this movement seek to develop psychological explanations of our intuitions that can help to assess their evidentiary value. Recent studies in cognitive epistemology develop this approach by first tracing intuitions back to the particular cognitive processes that give rise to them, and then identifying vitiating circumstances in which our intuitions can lead us astray (Nagel, 2010, e.g., Nagel, 2012b; Fischer, 2014; Alexander, Gonnerman, and Waterman, 2015; Gerken and Beebe, 2016; Fischer and Engelhardt, 2016; Fischer and Engelhardt, forthcoming)). Ultimately, the hope is to provide an 'epistemic profile' of the cognitive processes underlying philosophical intuitions that allows us to assess their evidential merit in specific types of philosophical inquiry.

A primary advantage of cognitive epistemology is that it facilitates a piecemeal approach to investigations of the nature and epistemology of intuitions. In line with our previous discussions, we suggest that this fine-grained approach to investigating the epistemology of intuitions is fitting to assess the viability of the expertise defence. In this sense, cognitive epistemology provides resources which can help to articulate the two claims constitutive of the expertise defence:

1. Philosophers have *philosophical expertise*—adequate training in philosophy leads to improved performance in some task(s).

2. Philosophical expertise *defends* philosophers from distorting factors that Critics claim undermines the practice of appealing to intuitions in philosophy.

If cognitive epistemology pinpoints biases and their deleterious effects in philosophical theorising, then presumably we can also use these findings to evaluate whether practices philosophers employ makes them less susceptible to these biases. That is, we can ask: do philosophers happen to do the sorts of things that shields them from the problematic effects of biases? We now argue that, for the application of one philosophical practice to one specific philosophical puzzle, that the answer is 'yes'.

[16] This term is coined by Fischer, (2014).

3.3.2 Philosophical Expertise and Ordinary Language Analysis

To demonstrate the relevance of recent empirical work in vindicating a narrow version of the expertise defence, we will look at the practices subsumed under the approach of Ordinary Language Analysis (OLA)—for helpful reviews, see Hansen, (2014) and Baz, (2016). Roughly put, OLA is the methodology which suggests that it can be philosophically fruitful to pay careful attention to how we use particular words in *everyday contexts*: the motivating idea is that this can help us to identify and resolve philosophical difficulties that arise from the mismatch in how these words are used in ordinary contexts, and in how philosophers use them in philosophical inquiry.[17] In what follows, we argue that there is empirical evidence showing that the practices of OLA, when applied within a particular line of inquiry, can give rise to a kind of philosophical expertise that shields philosophers from a specific type of bias.[18]

Before proceeding, we should note that we assume that when a philosopher uses OLA in attempting to solve a philosophical puzzle, the use of such techniques will be informed by previous philosophical training. Of course, philosophical training is not a *necessary* condition for deploying OLA—after all, this methodology presumably relies on general capacities we employ to make sense of discourse in everyday life. However, given that appeal to OLA in philosophical contexts requires careful attention to philosophical texts, together with the ability to effectively reconstruct argument and situate them among other philosophical theories, it is natural to expect that this skill will typically result from extensive training in philosophy.

One of the high-water marks of ordinary language analysis is Austin's *Sense and Sensibilia*. One of the primary targets here is Ayer's (1963) version of the so-called 'Argument from Illusion'—a historically influential argument that has shaped many debates in the philosophy of perception. This argument proceeds from descriptions of cases of non-veridical perception—i.e., cases where objects look or appear to have properties they do not actually possess. For instance:

(A) A stick that normally looks straight may appear bent when it is seen in water.

Ayer (ibid., p. 4) contends it is obviously implausible that a straight stick *becomes* bent once put in water, or that the stick can be simultaneously both crooked and straight. Therefore, if a stick can both appear to be straight in normal circumstances *and* look bent when put in water, then this suggests that in *at least* one of these cases we are not *directly aware* of the stick. Nevertheless, Ayer contends that even if the viewer is not directly aware of the stick, they must be aware of *something*. This 'something' is called a sense-datum, which is credited with the properties the stick appears to have.[19]

[17] For further discussion, see Hansen, (2014).

[18] Although OLA nowadays has something of a bad name, it is important to note that one can adopt the practices of OLA without subscribing to the more hefty methodological commitments that might have been by shared by its initial proponents. And indeed, the frequency with which considerations about linguistic data are brought to bear on assessments of philosophical theories shows that OLA is alive and well today. For a discussion, see Hansen, (ibid.); for a contemporary (qualified) defence of OLA, see Hansen and Chemla, (2015) and Hansen, (2017).

[19] Both Russell, (1912) and Broad, (1923) present the Argument from Illusion in ways very similar to Ayer. However, it is

One of Austin's central contentions is that proponents of the Argument from Illusion get us into a muddle because they neglect the subtle implications of the verbs 'looks', 'seems' and 'appears' when describing cases of non-veridical perception (Austin, 1962, pp. 33–43). As Austin aptly points out, proponents of the Argument from Illusion take themselves to use these *appearance-verbs* in a very distinctive way: specifically, they employ them to indicate *merely* how things look/appear/seem to a viewer 'there and then', without implying that viewers would be inclined to judge that things *really are* that way. Thus, they assume (quite plausibly) that most people who read descriptions like 'a straight stick half-submerged in water appears bent to S' will *not* infer that S would be inclined to *judge* that the stick really *is* bent. After all, most of us would know better than to judge that the stick *is* bent just because it appears to so be 'there and then'—we naturally assume that any other viewer would also know better.

However, as Austin skilfully teases out by way of a number of *linguistic intuitions*, clauses of the form "X seems F to S" and "X appears F to S" imply just that. That is, Austin proposes that these constructions *generally* convey that S would, lacking any countervailing information, judge that "X *is* F" (ibid., pp. 36–38). To illustrate, imagine I claim that 'the new bike lanes *appear* to be quite narrow', or 'the church *seems* to be gothic'. In these cases, we *ordinarily* take the speaker to endorse the content that comes after the appearance-verb. Of course, this *doxastic* implication can be *cancelled* by explicit contradiction: E.g., if we say 'the church seems gothic, but was actually constructed only 10 years ago', we take the relevant appearance-verb to indicate *only* how things seem to the viewer 'there and then', without implying that she judges things to be that way. However, absent such cancellations, we typically interpret these verbs in their doxastic sense, and infer that subjects would judge accordingly.

Why is any of this relevant for the Argument from Illusion? The reason is that, when describing cases of non-veridical perception, proponents of this argument intend to use appearance-verbs in a purely *phenomenal* sense: that is, they take these verbs to indicate only how things seem to a subject from that perspective, without implying anything about their inclination to judge that things really are as they appear, look, or seem to them 'there and then'. However, as Austin points out, uncancelled uses of appearance-verbs ordinarily have just this implication; moreover, this suggests that uncancelled use of these verbs when describing non-veridical perception cases for the Argument from Illusion might not suffice to indicate that these verbs are meant in their purely phenomenal sense (ibid., pp. 36–41). Therefore, as Austin is at pains to do throughout, it is very important to make explicit exactly what is intended by the use of certain appearance-verbs, and what is already known by the viewer. For instance, one needs to be reminded of the situation one is in when it is claimed that a stick half-submerged in water *appears* bent, and how this bears on the reading of the relevant appearance-verb; after all, it is not implausible to think that in some cases one could in fact suspect that the stick *really is* bent—say, when one did not previously know that the stick was actually straight. Thus, Austin's use of OLA calls attention to the fact that, without due caution in our philosophical discussions, certain uses of appearance-verbs can exert potentially confusing influences on our judgements.

noteworthy that contemporary reconstructions of this argument differ considerably from these formulations (see, e.g., Smith, 2002).

So much for Austin's application of Ordinary Language Analysis to the Argument from Illusion; why think that this has any relevance to a successful expertise defence? Here's why: in a number of recent papers, Fischer and colleagues (Fischer, Engelhardt, and Herbelot, 2015; Fischer and Engelhardt, 2016; Fischer and Engelhardt, forthcoming) provide empirical evidence that supports Austin's suspicions that misuse of appearance-verbs in descriptions of cases of non-veridical perception gives rise to significant philosophical difficulties. In what follows, we first present details of this empirical research, and then build on their findings to argue that Austin's use of OLA in analysing the Argument from Illusion is a clear display of significant philosophical expertise that does lend support to the expertise defence.

The empirical research by Fischer et al. (Fischer, Engelhardt, and Herbelot, 2015; Fischer and Engelhardt, 2016; Fischer and Engelhardt, forthcoming) is a prime example of a project in cognitive epistemology: it aims to provide a psychological explanation of the intuitions driving the Argument from Illusion, and then builds on this explanation to assess their evidentiary value. The psychological explanation Fischer et al. provide builds on a body of work in cognitive psychology on the process called *stereotype enrichment*. This is a cognitive process that is key to our competence as speaker/hearers of a language, and which plays a central role in both language production and comprehension (Hare et al., 2009; Ferretti, McRae, and Hatherell, 2001; Atlas and Levinson, 1981). To illustrate the operations of this process, consider the following claim:

(Tomato) "She looked at the vibrant colour of the tomato"

Most people who read the claim above infer that I was looking at the colour red—rather than green, orange, or yellow. Why do people immediately leap to this conclusion, since the tomato could have been of any one of these other colours? One well-accepted explanation is that, for most of us, tomatoes have a strong *stereotypical* association with the colour red; this strong association is what facilitates the inference to the conclusion that the viewer faces a particular colour (ibid.). In this sense, we "enrich" the utterance with our stored information about the colours that are stereotypically associated with tomatoes.

Stereotype enrichment is a generally reliable process for arriving at accurate conclusions about the world.[20] For instance, most of the tomatoes we encounter in our daily lives are in fact red. However, there are obviously cases where such stereotype enrichments can lead us astray. The following riddle is helpful to illustrate (Giora, 2003, p. 13):

(S) A boy and his father get into a severe car accident. The father dies immediately, and the boy is taken to hospital. In the emergency room, the surgeon looks at the boy and says 'I can't operate on him. He's my son.' How is this possible?

[20] The explanation for this is that stereotypical associations are encoded in our semantic memory, which evolves from frequent activation by repeated encounters with certain stimuli in the world (Neely, 2012). Given that the strength of a link in semantic memory is directly correlated to the observed co-occurrence of frequencies in the subject's environment, semantic memory will then tend to encode accurate stereotypes (e.g., most tomatoes are red). Of course, stereotypes are not always correct, and may be encoded by misleading information in the world—e.g., biased media and uncooperative testimony.

Many people, when first exposed to this question, are stumped by it. A plausible explanation for why people tend to be puzzled is the fact that the term 'surgeon' has a strong gender stereotype. This strong stereotypical association immediately prompts the leap to the conclusion that the surgeon must be the father—which, in turn, gives rise to the confusion ('But the father has died!'). This example illustrates how, even under situations where we would expect the contextual information to trigger the non-stereotypical meaning (viz., the use of 'surgeon' to pick out the boy's mother), we may still draw inferences that are 'enriched' by stereotypical associations.

In recent work, Fischer et al. (Fischer, Engelhardt, and Herbelot, 2015; Fischer and Engelhardt, 2016) argue that a similar kind of bias drives the intuitions underlying the Argument from Illusion. One of the central lines of empirical support for this claim is the evidence that *all* appearance-verbs (seems, appears, and looks) have very strong stereotypical association with certain doxastic properties. In more detail, Fischer et al. provide substantial empirical evidence—by way of a variety of experimental methods—for the following proposal:[21]

(H) The assertion 'x looks/appears/seems F to S' stereotypically implies that *S is inclined to judge that x is F*

The central upshot of **(H)** is that people will *ordinarily* interpret *all* appearance-verbs in their *doxastic sense*—i.e., to imply specific doxastic properties about the viewer. This in itself may sound rather unsurprising; after all, the work of Austin already suggests as much: we tend to interpret clauses of the form "x seems F to S" and "x appears F to S" in this dominant doxastic sense, unless this meaning is explicitly ruled out.[22] However, the findings from Fischer et al. (Fischer, Engelhardt, and Herbelot, 2015; Fischer and Engelhardt, 2016; Fischer and Engelhardt, forthcoming) build on Austin's suggestion in an important way: their findings show that these stereotypical associations are so strong that people will tend to interpret these verbs in their doxastic sense *even in situations which invite the non-doxastic (phenomenal) interpretation of these verbs*.

This finding is surprising; because, as mentioned above, proponents of the Argument from Illusion take it that people will interpret these verbs in their purely *phenomenal* sense when they read descriptions of non-veridical cases of perception. That is, they assume (quite plausibly) that in reading descriptions of

[21] These findings were obtained from two distinct experimental methods. The first is a distributional-semantic analysis of these verbs within a vast linguistic corpus, which suggests there to be strong links between these verbs and the relevant doxastic properties. The second was the use of a forced-choice task where participants were asked to rate the plausibility of items where the complement of the appearance-verb was inconsistent with doxastic inferences. Results showed 'look' and 'appear' to be strongly associated with doxastic features, and 'seem' to be *even more* strongly associated with such features. Notably, this last method can also be combined with experimental techniques of pupillometry, which tracks increase in levels of cognitive effort by measuring the dilation of participants' pupils. This method has been fruitful in investigation of perception-verbs ('see' and 'aware'), where results obtained from previous forced-choice plausibility-ranking tasks with such verbs (Fischer and Engelhardt, 2017) were replicated in studies employing techniques of pupillometry (Fischer and Engelhardt, forthcoming).

[22] Interestingly, Fischer et al's work show that *all* appearance-verbs (including clauses of the form "x looks F to S", whose doxastic uses Austin (1962, p. 36) dismisses) have strong stereotypical associations with certain doxastic properties.

cases such as 'a large boat viewed from afar *appears* small', people will resist inferring that the viewer in these cases is inclined to judge that the object viewed *is actually* small. Again, this assumption is well-motivated: after all, most of us know very well that a large boat does not *actually* possess the property of smallness because it is viewed from afar—and we take it that other people would also not be tempted to so judge. However, the findings from Fischer et al. show just that: people will indeed tend to interpret these verbs in their doxastic sense when reading these kinds of cases—thus implying that in this case the viewer would be inclined to judge that the object does have the properties it looks/appears/seems to have. This in itself calls for an explanation: given the plausible assumption that people would interpret these verbs in their phenomenal sense when reading the kinds of cases that motivate the Argument from Illusion, why is it that they still take them in their doxastic sense?

Relying on a large body of work from research in cognitive psychology, Fischer and colleagues propose there is a rather natural explanation for this tendency. A clue to the answer is found in explanation of the small riddle of the surgeon (**S**) presented above. Recall, the explanation for why it can be so difficult to solve this riddle is that the term 'surgeon' has a very strong gender stereotype with 'male'. Indeed, as the riddle helpfully illustrates, mere mention of the term will immediately prompt the stereotype-driven inference that the surgeon is the father—even when the context indicates that this conclusion is not possible. A similar kind of explanation can account for people's tendency to interpret appearance-verbs in their doxastic sense when reading descriptions of cases of non-veridical perception. More precisely, the strong stereotypical associations between appearance-verbs and the relevant doxastic properties immediately drives the stereotype-driven inference that viewers are inclined to judge that things are as they look/appear/seem to them. Moreover, like in the riddle of the surgeon, this inference is activated even when contextual information indicates otherwise—e.g., even in the light of the explicit information that the stick that looks bent *is* straight.

Fischer et al. (Fischer, Engelhardt, and Herbelot, 2015; Fischer and Engelhardt, 2016) argue that this provides a psychological explanation for why proponents (and readers) of the Argument from Illusion effortlessly leap from descriptions of cases of non-veridical perception to the negative conclusion that in such cases we are not aware of objects themselves. Very roughly put, the reason why this conclusion seems to follow so naturally is that when people read, e.g., 'a straight stick half-submerged in water looks bent', they then immediately infer that the viewer is inclined to judge that the object viewed *is* bent. However, it is explicitly clear from the description of the case that the stick is straight, rather than bent. This conflict between the stereotype-driven inference (that the object viewed is judged to be bent) and the explicit description of the case (where the stick is described as being straight) then leads to the conclusion that the viewer does not know what object she is viewing, and is not aware of that object—but of something else.

The psychological explanation above reveals how the Argument from Illusion gains its intuitive force by trading on a confusion between the *phenomenal* and *doxastic* senses of the relevant appearance-verbs. This then suggests that the Argument from Illusion is underwritten by an incisive kind of cognitive bias and that the intuitions driving this argument are epistemically worthless. The particular bias in question is

traceable to an identifiable set of vitiating circumstances: specifically, this bias arises because proponents of the Argument from Illusion make use appearance-verbs in their less salient *phenomenal* sense when describing cases of non-veridical perception, without making it explicit that this use deviates from the (dominant) doxastic sense of these verbs. The unmarked use of this less salient technical sense leads competent speakers to mistakenly draw inferences in line with the ordinary (doxastic) sense of this verb, which conflict with the inferences that are natural to draw with respect to its phenomenal sense. Moreover, it is precisely the confusion between the doxastic and phenomenal senses of the relevant verbs that prompts the conflicting intuitions driving the Argument from Illusion. Given that such linguistic confusions should have no bearing whatsoever on philosophical investigations of the nature of perception, then it is plausible to conclude that these intuitions are epistemically defective.

Let us now return to the discussion of the expertise defence. We propose that the empirical work of Fischer and colleagues reveals that Austin's use of OLA amounts to a form of philosophical expertise that does suffice to shield philosophers from certain biases, and therefore supports an instance of the expertise defence. As discussed above, one of Austin's central contentions was that proponent of the Argument from Illusion get us into a muddle when they use appearance-verbs without due attention to their subtle pragmatic implications. By paying attention to these implications, we can avoid the misleading influences they can exert on our judgements. To support his claim, Austin employed a Critical use of the vignettes motivating the Argument from Illusions. As the empirical work of Fischer et al. demonstrate, Austin's Critical examination of these cases is correct: the confusion Austin identifies is indeed the source of what we take to be an uncontroversial instance of a bias.[23] The bias in question stems from the operation of a generally reliable cognitive process known as stereotype enrichment which, due to the vitiating circumstances arising from misuse to the subtle implication of appearance-verbs, gives rise to intuitions that have no evidentiary value. In this way, the empirical evidence reviewed above shows that a specific form of philosophical training (in OLA) can lead to improved performance in particular practices surrounding the use of intuitions (using thought-vignettes in a Critical way to uncover *and suppress* a bias stemming from a specific cognitive process) for the purposes of improved philosophical inquiry into the nature of perception.

Does this mean that ordinary language analysis suffices to defeat biases across the board? Almost certainly not. As we have attempted to emphasise throughout this chapter, more care is required in attributions of philosophical expertise, and relating it to the expertise defence. We have not provided evidence that ordinary language philosophy can help to avoid problematic effects of other biases, or that its practices are truth-conducive in other instances. But, this is not an unwelcome upshot of our view; as we have suggested, this piecemeal approach is required to do justice to the complexity of philosophical inquiry, and to provide a realistic evaluation of the expertise defence. It remains to be seen whether the expertise defence can be upheld for other practices, in other lines of philosophical inquiry. In advance of the careful

[23] It is worth noting that this discussion suggests a further locus for philosophical expertise: framing verbal descriptions of cases in such a way that can avoid engaging biases arising from linguistic confusions in the first place. This form of expertise can be enabled by using empirical research in psycholinguistics, which reveals the precise ways in that these linguistic confusions can lead us astray. E.g. see Fischer and Engelhardt, (forthcoming).

empirical work that is required to identify the processes driving cognitive biases, and considering how these intersect with our current philosophical practices, we recommend an attitude of studied agnosticism.

3.4 Conclusion

We have been exploring the viability of the expertise defence as a response to the experimentalist challenge. By critiquing extant positions, we found that both parties in this debate operate with an inadequate conception of philosophical expertise. In responding to these worries, we have offered a novel account of philosophical expertise that is sensitive to the variety of different practices and methods that comprise philosophical inquiry. This account better reflects the complex and variable relationship between philosophical expertise and experimentally-motivated concerns about intuitions. The upshot of our account is that we cannot vindicate or reject the expertise defence all at once. Rather, we must adopt a nuanced and localised approach that attends to the details of precisely what philosophical methods are being used, and whether these methods suffice to shield philosophers from specific distorting effects that skew our intuitions. And finally, we illustrated how one might take up our novel approach by appealing to recent work in cognitive epistemology, which we propose supports a local instance of the expertise defence. Our arguments show that much work remains to be done to determine whether appealing to philosophical expertise is a viable route. Moreover, these arguments indicate the importance that experimental approaches will have in this project.

Chapter Four

Who's Afraid of Cognitive Diversity?

In previous chapters, I have argued against prominent formulations of the radical version of the experimentalist challenge—that is, the proposal that empirical studies reveal that philosophers should refrain from appealing to intuitions. The arguments I developed aimed to show that the *current* body of empirical evidence does not warrant such a strong conclusion.[1] However, these arguments do not rule out the radical experimentalist challenge. For instance, they still leave it open that further studies could amass sufficient empirical evidence to show that all intuitions are prone to some deleterious effect or other. Some Critics have claimed that the upcoming international research efforts testing for cross-cultural variation in people's intuitions are a promising step in this direction.[2] Their expectation is that these efforts will reveal significant demographic variation in how people respond to philosophically relevant thought-experiments. And, given the seemingly compelling claim that intuitions that are found to be sensitive to epistemically irrelevant factors are unsuitable to be used in philosophy, they suggest that these findings would provide strong support for a radical version of the experimentalist challenge.

In this chapter, I argue that there are good reasons to resist these radical claims by the Critics. My main contention is that intuitions found to be demography-specific *need not* be deemed unsuitable for use in philosophical inquiry. In the first section, I further explain the proposal that findings of demographic effects show that philosophical views that rely on that intuition are problematic (sec. 4.1). I then provide a novel analysis of this challenge by outlining the epistemological considerations driving its main claims. Building on this analysis, I tease apart two distinct formulations of this challenge (sec. 4.2). In the subsequent sections, I explain why there are good reasons to resist both of these formulations (secs. 4.3 & 4.4). In effect, the arguments in this chapter demonstrate that evidence of cross-cultural variability is not as damning as Critics would have us believe.

[1] As this conclusion either relies on a problematic inductive claim about the unreliability of all intuitions (chap. 2), on claims about their overall hopelessness (chap 2), or on the idea that the intuitions of philosophers are on a par with those of non-philosophers (chap 3).

[2] For details, see //`http://www.thegeographyofphilosophy.com/`//./

4.1 The Challenge from Cognitive Diversity

Before proceeding, it will be helpful to make two preliminary points for the discussion that follows. First, I denominate the body of evidence showing people's intuitions about philosophical thought-experiments to vary along *demographic* lines as 'findings of cognitive diversity'. Second, my use of this expression is meant to capture *all* vectors of such demographic variation—e.g., across cultural divides, socio-economic levels, and degrees of philosophical training. There are of course a number of important methodological distinctions to be drawn between these many vectors; however, for current dialectical purposes such distinctions will not be important (and so I ignore them in what follows). With these considerations in mind, let us now turn to the methodological challenge that I will be examining in this chapter. To give this an initial gloss, consider the following case:

> Suppose that an essential part of a philosopher S's argument for a theory R in epistemology is that many people in S's environment report having an intuitive judgement J in response to a philosophical thought-experiment T. However, suppose that empirical findings reveal that people from different cultures (other than S's and of her interlocutors) tend to have a *diverging* intuitive judgement J* with respect to T—and that J* supports the rival epistemological theory R*. S acknowledges these diverging intuitive judgements about T from people of other cultures; yet, she claims that, because her intuitive judgement J seems *self-evidently* and *intuitively* correct, and because it is shared with people in her community, then this provides reasons to believe R.

Most people agree that there is something amiss with S's appeal to J in support of her theory R. In particular, it seems dubious for S to assume that belonging to her demographic group puts her in a privileged epistemic position to judge on philosophical matters—indeed, it is a sensible assumption that features of demography should play no role *whatsoever* in philosophical debates. Thus, S's appeal to her own intuitive judgement J (at the cost of ignoring those of people from other cultures) strikes many as *dogmatic*.[3] Furthermore, *Critics* argue that we can extend the verdict above to *all* cases where intuitive judgements are found to vary along demographic lines (for a review: Stich and Tobia, 2016). In a forceful passage that is representative of this line of thought, Stephen Stich suggests that

> unless one is inclined toward chauvinism or xenophobia in matters epistemic, it is hard to see why one would much care that [an intuition] one was thinking of invoking (or renouncing) accords with the set of evaluative notions that prevail in the society into which one happened to be born.
>
> (Stich, 1990, p. 94)

[3] I am restricting my use of the term 'dogmatic' to a negative evaluation regarding the unreasonableness of one's propositional attitudes. For a defence of the potential reasonableness of dogmatic beliefs, see Kelly (2011).

And similarly, Edouard Machery says:

> Philosophers do not realize how contingent on a particular historical, cultural, and social position their formulation of the fundamental philosophical questions and the answers they happen to consider are, and they ignore many of the options available to them when addressing these philosophical questions, limiting themselves to those few options that, for accidental reasons, happen to be deemed worth considering.
>
> (Machery, 2017, p. 145)

Building on these and other similar considerations, Critics have argued that evidence of cognitive diversity about any thought-experiment T gives us *decisive* reasons to think that intuitive judgements about T should not pull any weight in philosophical discussions. More schematically:

Challenge from Cognitive Diversity (CCD)

P1 If there is evidence of cognitive diversity with respect to a thought-experiment T, then appealing to an intuitive judgement about T in philosophical inquiry is *dogmatic*.

P2 If a judgement is dogmatic, then that judgement does not provide warrant for philosophical positions or claims.

C Therefore, if there is cognitive diversity about a thought-experiment T, then appeals to intuitive judgements about T fail to provide warrant for philosophical position or claims.

Critics regard the CCD to be compelling. Many have agreed that it is indeed hard to shake off the idea that if intuitions are *actually* demography-specific, then it would be ultimately illicit to rely on them in philosophical inquiry. However, of the very few studies reporting demographic variation in intuitive judgements (for a review: ibid., ch. 2), many have either failed to replicate, or have been shown to suffer from severe methodological flaws (see, e.g., Adleberg, Thompson, and Nahmias, 2015; Kim and Yuan, 2015; Seyedsayamdost, 2015a; Seyedsayamdost, 2015b; Freitas et al., 2018).[4] So, at the moment, the threat of the CCD does not seem to loom large.

However, this reveals in what way the nascent international studies might prove important. The methodology these studies adopt aims to correct for the blatant shortcomings of previous empirical investigations. First, they greatly widen the scope of extant findings, testing a great number of other intuitions. This could then potentially considerably expand the (until now) very limited body of evidence of cognitive diversity. And second, these studies aim to correct for many of the glaring methodological errors of previous experiments—employing more refined and appropriate experimental techniques for the purposes of testing people's intuitions. Thus, Critics anticipate that the upcoming studies could have

[4] For a notable exception to these failures of replication and methodology, see Machery et al. (2004).

revolutionary implications. For, given that the CCD is indeed compelling, then the prospect that these studies may uncover many novel findings of cognitive diversity could then undermine a large swathe of our firmly-held philosophical views.

For the remainder of this chapter, I argue that although *seemingly* compelling, the CCD can be shown to rest on highly controversial assumptions. My argumentative strategy is composed of two steps. In the first step, I underscore the methodological and epistemological considerations driving the first premise of the CCD. In the second step, I explain why there are very good reasons to question that these considerations successfully motivate the conclusion of this challenge.

4.2 A Closer Look at the CCD

In this section, I give a more detailed analysis of the motivations for the CCD. My main purpose is to identify the concerns and epistemological principles driving the first premise of this challenge—viz., the claim that appealing to intuitive judgements found to vary along demographic lines is ultimately dogmatic. As we will see, this detailed analysis provides useful resources with which to conduct a better assessment of the shortcomings of the CCD. As a first step to developing the arguments in this section, I now briefly describe recent debates about the phenomenon of *contingency* anxiety: that is, the sense of unease that ensues after discovering that one holds a belief due to influences of *arbitrary factors* (see, e.g., White, 2010; Schoenfield, 2014; Mogensen, 2016; Vavova, 2018).[5]

4.2.1 Contingency Anxiety

Discussions about the phenomenon of contingency anxiety are primarily driven by the observation that demographic factors seem to determine what beliefs we come to hold. For instance, it is remarkable that people often have religious, political, and moral beliefs that align *very closely* to those of other people in their community. Now, most of us would like to think that our beliefs are due to careful consideration of the evidence available to us; but, if arbitrary factors of our demography determine what we believe on such important issues, then it seems that we believe as we do for reasons that are far-removed from their truth. This then gives rise to a sense of unease about our beliefs: for, the realisation that you could have easily believed otherwise (say, if you had been born elsewhere, or gone to a different school) calls into question the grounds for those beliefs.

There are two issues that arise in connection with debates about contingency anxiety that are worth flagging up here. The first is definitional: for the purposes of this chapter, I take a factor C to be *arbitrary* with respect to a subject S's belief that p if S's knowledge that C obtains does not give S reasons to believe that p.[6] For instance, learning that water is H20 gives me absolutely no reason to believe that the word 'apple' has 5 letters; and in this sense, my belief that water is H20 is arbitrary relative to my belief that 'apple' has 5 letters.

[5] I am certainly not the first person to call attention to the relevance of debates about contingency anxiety with regards to the CCD. But, the particular framing I give in this section is indeed novel.

[6] This definition is modelled on the one provided by Andreas Mogensen (2016).

A second issue is that, clearly, not all influences of arbitrary factors on belief are problematic. For instance, suppose that in an attempt to combat my insomnia, I go out for a walk at 3 am. As it so happens, during my walk I discover that the corner shop is open at 3 a.m. Now, it is clear that the fact I have insomnia does not itself provide reasons to hold beliefs about the corner shop's opening hours. Thus, in line with the above definition, the fact that my insomnia in some way caused me to believe that the shop is open at 3 am is an arbitrary factor in my belief. Yet, unlike the cases of influence from demographic factors on belief, the fact that my insomnia caused my belief about the corner shop's opening hours seems untroubling.[7] This then raises the following crucial question: as opposed to the unproblematic influence of my insomnia (an arbitrary factor) on my belief about corner shop's opening hours, what is it about arbitrary *demographic* factors on belief that gives rise to contingency anxiety? Focusing on this question will prove to be useful for the purposes of this chapter: as we will see, answering it will help us to underscore the motivations that drive the first premise of the CCD.

We can gain further understanding of the phenomenon of contingency anxiety with respect to arbitrary demographic factors by considering the following case (Schoenfield, 2014, p. 204):

> **ZAPPER**: You start out having no idea whether *p*. You then look at some evidence which either entails *p* or entails ~p. (Suppose that there is only one rational response to this evidence: if it entails *p*, you should believe *p*, and if it entails ~p, you should believe ~p). You reason and conclude that *p*. You then learn that a scientist flipped a coin to determine what you will believe. She decided that, if the coin lands heads, she will zap your brain so that it seems to you that there is a good argument for *p*, and if it lands tails, she will zap your brain so that it seems to you that there is a good argument for ~p.

Presumably, learning about your predicament in the case of Zapper gives you strong reasons to doubt your judgement on whether *p*. After all, it is very likely that your judgement on whether *p* will turn out to be *incorrect* precisely because the scientist determined what you will judge about *p*. There are (at least) two different ways of fleshing out this claim. The first is driven by the idea that being zapped by the scientist causes you to not respond *rationally* to the evidence; the second centres instead on the claim that being zapped causes you to judge *wrongly* on whether *p*. In this sense, the case of Zapper gives rise to two distinct hypotheses:

1. Your analysis of whether *p* was irrational.[8]
2. Your analysis of whether *p* was false.

(To a first approximation these two hypotheses might seem inextricable. After all, we often think that being rational leads to false beliefs. But, note that these two dimensions of epistemic evaluation do indeed seem to come apart in some cases. To illustrate, consider that it could might as well be the case that even

[7] This example is loosely based on Mogensen's (ibid.).

[8] This is restricted to concerns about irrationality that arise before learning of your predicament.

if you by luck were right in your judgement that *p*, it is nevertheless possible that the scientist's influence made it that you were irrational in arriving at that judgement. For instance, being zapped might have lead you to just problematically ignore part of the evidence. In this sense, distinguishing between the above two hypotheses helps to hone in on the different kinds of considerations pertinent to an analysis of Zapper.)[9]

Focusing on the case of Zapper is helpful to elucidate what it is about learning of the influence of arbitrary *demographic* factors that gives rise to contingency anxiety. That is because the influence of our demographic factors on beliefs is, in effect, very much like that of the scientist on your judgement in Zapper. For, as in the case of the scientist's influence on your judgement that *p*, the influences of your demographic group presumably determine what you believe for reasons that are possibly removed from the truth of the matter. Thus, as in Zapper, the influence of our demographic factors make it likely that our beliefs may turn out to be incorrect. Again, we can flesh this out in at least two ways. First, the influence of your demographic factors can be taken to make you irrational in how you respond to the evidence available to you. For instance, it could be that your demographic factors lead you to irrationally favour some pieces of evidence whilst neglecting others. Alternatively, the influence of your demographic factors might lead you to hold false beliefs (even if you evaluated the evidence in a rational way). In this sense, it is somewhat clear why it seems appropriate to feel uneasy about those beliefs upon learning of the influence of arbitrary demographic factors on them. And importantly, note that a similar verdict is not applicable to the case where my insomnia influences my beliefs about the corner shop's opening hours. After all, whether I have insomnia or not does not in any way speak to the potential irrationality or falsity of such beliefs. This then pinpoints just why this case (albeit picking out a case where arbitrary factors influence belief) does not give rise to contingency anxiety (for a discussion: Vavova, 2018).

There is one final observation that is worth flagging up at this point. As discussed above, most people would agree that learning of your predicament in the case of Zapper would give you strong reason to reduce your confidence in your judgement on whether *p*. Now, consider the following line of reasoning in response to this verdict: "I recognise that the scientist's influence on my belief means that I probably made a mistake in evaluating the evidence on whether *p*. But, after going through my reasoning again, I remain convinced that *p* is the answer. My reasoning is indeed sound. So, it seems clear that I should conclude that *p* is true". What is problematic about this reply is that it *dogmatically* resorts to your own reasoning about whether *p* to argue against the hypotheses that this same line of reasoning was either false or irrational. After all, your reasoning about *p* is precisely what is called into question by the discover of the scientist's influence. On a similar note, it is clear that the following reply would be ineffective in cases where one learns of arbitrary demographic influences on belief: "I recognise that the influence of my demographic features on belief make it likely that my belief is mistaken. But, having gone through the reasons for my belief, I find them to be compelling. So, it seems clear that my belief is true." Again, the problem is that this reply makes use of precisely the line of reasoning that is put in question after finding out about the influence of demographic factors. So, it seems that in attempting to use your own belief in

[9] This perspicuous analysis of the case is due to Schoenfield (2014).

this case, it is clear you are holding a *dogmatic* belief.

4.2.2 The CCD and Contingency Anxiety

Let us now return to the CCD. I will argue that the above discussions about the phenomenon of contingency anxiety pinpoint the concerns driving the first premise of this challenge—viz., the claim that appealing to intuitive judgements found to vary along demographic lines is ultimately dogmatic. Building on these considerations, I distinguish between two ways of formulating the CCD—which, as we will see in the next sections, prove useful for providing a more adequate and thorough assessment of this challenge.

As a first step to developing the arguments in this section, note that findings of cognitive diversity plausibly give rise to contingency anxiety.[10] For, what these findings indicate is that people with distinct demographic features will have different intuitive judgements. In this sense, these findings reveal that arbitrary features of our demography may be that which ultimately determine our intuitive judgements about philosophical thought-experiments. Thus, as in other cases of arbitrary demographic factors on belief, findings of cognitive diversity in intuitive judgements seem to call our judgements into question, since they suggest that if things had been different (e.g., if we had been born elsewhere, or gone to a different school) then we would have had different intuitive judgements.[11]

Framing evidence of cognitive diversity in this way is useful for current purposes. In particular, these observations underscore an important parallel between cases of cognitive diversity and the kind of scenario described in Zapper: in both cases, the influence of an arbitrary factor on judgement provides strong reason to think that judgements are very likely to be *mistaken*. Thus, in line with our above analysis of Zapper, cases of cognitive diversity can also be taken to raise the following two hypotheses:

1. Your intuitive judgement of the philosophical thought-experiment was irrational.[12]
2. Your intuitive judgement of the philosophical thought-experiment was false.

Now, note that as with the cases of Zapper and of arbitrary demographic factors on belief, there does seem to be something terribly dubious about responding to these hypotheses as follows: "I recognise that in cases such as these, there is a significant chance that I could have made a mistake in my evaluation of the case. But, going over the case, I remain convinced that I have responded to it correctly: my intuitive judgements just seems evidently right to me. So, my intuitive judgements is correct." Once more, what is troubling about this response is that it attempts to dismiss the possibility that your intuitive judgement is

[10] For an example of this kind of reading of the findings of cognitive diversity, see Knobe and Nichols (2008, p. 11).

[11] Although evidence of cognitive diversity will thus uncover instances of *disagreement* about philosophical cases, note that the kind of disagreement they reveal is *particularly* troubling. To appreciate this claim, note that there is unlikely to be any philosophical intuition that is universally shared. But, such disagreements do not *uncontroversially* give sufficient reason to think that it would be dogmatic to appeal to disputed intuitive judgements in philosophical inquiry. By contrast, the fact that disagreements about intuitive judgements could be traced back to arbitrary demographic factors raises a pernicious challenge to many philosophical positions and claims. Thus, even if we take evidence of cognitive diversity to uncover instances of disagreement, this is a particular brand of disagreement that seems more troubling for the prospects of using intuitive judgements in support for philosophical positions and claims.

[12] Again, I am restricting this to concerns about irrationality that arise *before* learning of the evidence of cognitive diversity.

either irrational and false by relying precisely on the intuitive judgement that is called into question by the evidence of cognitive diversity. Thus, similar to cases discussed in the previous section, it seems clear that maintaining that judgement in light of that evidence (and of the hypotheses of irrationality and falsity that it raises) will be ultimately *dogmatic*.

The above considerations considerably help explain why the charge of *dogmatism* seems appropriate in light of evidence of cognitive diversity. In more detail, I propose that findings of demographic variation in our intuitions raise the hypothesis that our intuitive judgements are likely to be either irrational or false. And, as explained above, attempts to appeal to our intuitive judgements in light of these possibilities is dogmatic. Thus, I propose that what motivates the charge of dogmatism in the CCD are concerns arising from the hypotheses of irrationality or falsity. Building on these considerations, we can then distinguish two versions of the CCD. The first is driven by threat of irrationality:

Challenge of Irrationality from Cognitive Diversity (CICD)

P1 Findings of cognitive diversity about a thought-experiment T provide strong undefeated evidence that intuitive judgements about T are irrational.

P2 If there is undefeated evidence that a judgement is irrational, then appeals to that judgement fail to provide warrant for philosophical positions and claims.

C Therefore, if there is cognitive diversity about a thought-experiment T, then appeals to intuitive judgements about T fail to provide warrant for philosophical position and claims.

By contrast, a second version CCD is driven by the threat of falsity arising from evidence of cognitive diversity:

Challenge of Falsity from Cognitive Diversity (CFCD)

P1 If there is cognitive diversity about a thought-experiment T, then this provides strong undefeated evidence that intuitive judgements about T are false.

P2 If there is undefeated evidence that a judgement is false, then appeals to that judgement fail to provide warrant for philosophical positions and claims.

C Therefore, if there is cognitive diversity about a thought-experiment T, then appeals to intuitive judgements about T fail to provide warrant for philosophical position and claims.

This ends the first step of the argumentative strategy I adopt in this chapter. So far, I identified the epistemological and methodological considerations driving the charge of dogmatism in the first premise of CCD. Building on recent debates about contingency anxiety, I argued that evidence of cognitive diversity about a thought-experiment T provides undefeated evidence that intuitive judgements about T are either

irrational or false. Building on this analysis, I distinguished between two versions of the CCD: the first is driven by the concern of irrationality arising from evidence of cognitive diversity (CICD) and the second is driven by concerns of falsity (CFCD). In the next two sections, I develop the second part of the argumentative strategy of the chapter. My aim there will to be show how there are good reasons to resist both the CICD and the CFCD insofar as they require a defence of controversial positions.

4.3 Against the CICD

In this section, I argue against the first premise of the CICD: i.e., the proposal that evidence of cognitive diversity about a thought-experiment T provides strong undefeated evidence that intuitive judgements about T are *irrational*. As a first step to developing this argument, I first briefly explain the motivations for the view known as 'Permissivism': that is, the view according to which a body of evidence can rationalise more than one doxastic attitude.[13] Later in the section, I argue that an articulation of the CICD will require maintaining that Permissivism is false—i.e., the view known as 'Uniqueness'. In this light I propose that the CICD is highly controversial insofar as debates between Permissivism and Impermissivism are one of the most disputed and intricate debates in contemporary epistemology. I then consider and reject a reply available to Critics against this proposal.

4.3.1 Permissivism

As already stated, Permissivism can be roughly formulated as the view that, possibly, a body of evidence can rationalise more than one doxastic attitude. Typically, arguments for Permissivism are motivated by considering cases where people might plausibly be said to be rational in their disagreement about what a body of evidence supports. Speaking to this issue, Gideon Rosen says that

> [i]t should be obvious that reasonable people can disagree, even when confronted with a single body of evidence [...] Palaeontologists disagree about what killed the dinosaurs. And while it is possible that most of the parties to this dispute are irrational, this need not be the case. To the contrary, it would appear to be a fact of epistemic life that a careful review of the evidence does not guarantee consensus even among thoughtful and otherwise rational investigators.
>
> (Rosen, 2001, p. 71)

Permissivists suggest that focusing on other disputes over difficult issues in the sciences, law, and philosophy will motivate a similar conclusion (for a review: Kopec and Titelbaum, 2016). For instance, consider the thorny disputes between Lewis and Stalnaker (and between their respective students) on how to best flesh out the notion of possible worlds. Again, even though Stalnaker and Lewis disagree on what

[13] Notably, there are many ways of characterising the Permissivist position. I have here adopted the formulation in terms of how many *doxastic attitudes* a body of evidence can rationalise; but, nothing hangs on this choice (for a detailed analysis: Titelbaum & Kopec, ms.).

the available evidence supports, still there does indeed seem to be some temptation to think that *both* could be rational in maintaining their respective diverging positions.

Now, although these brief considerations give initial reasons to think there are indeed such Permissive cases, it will here be useful to flesh out in a bit more detail just *why* the above examples can be said to motivate this conclusion. For this aim, it will be useful to consider a second set of arguments for Permissivism.

As many have pointed out, the denial of Permissivism (i.e., the view known as Uniqueness) commits us to a conception of evidential support in terms of a two-place relation of the form 'E does/doesn't support P' (Douven, 2009; Kelly, 2013; Kopec and Titelbaum, 2016). On this view, whether a body of evidence E provides epistemic support for a proposition P is defined *entirely* in terms of the body of evidence and the relevant proposition—nothing else needs to be factored in. However, as many have taken pains to show, conceiving of facts of evidential support in this way is either controversial (Douven, 2009; Kelly, 2013; Ballantyne and Coffman, 2011), or simply untenable (Decker, 2012; Titelbaum and Kopec, ms.). Delving into the finer details of these arguments is beyond the scope or aims of this chapter. We need only point out that common to all these proposals is the idea is that the two-place relation view of evidential support fails for the same reason that objectivist theories of confirmation theory do (e.g., Carnap's notion of logical probability, Hempel's positive instance account): namely, they are unable to specify the *grounds* which make a proposition justified with respect to a body of evidence.

In response to the above considerations, many have suggested that facts of evidential support should be taken as three-place relations, where the third relatum specifies what these grounds are. There are many explanations for filling out the third relatum in evidential support relations—e.g., an interpretation (Decker, 2012), prior probability functions (Meacham, 2013), an agent's epistemic goals (Kelly, 2013), or their epistemic standards (Schoenfield, 2014). Despite many differences in these proposals, I follow Titelbaum and Kopec (n.d.) in defining all such views as committed to the idea that this third relatum specifies a (roughly defined) *method* for analysing a body of evidence in order to derive conclusions. Thus, on a Bayesian framework, these methods will specify the subject's priors; alternatively, we can think of methods as a set of epistemic standards expressed as norms of the form 'if E, then believe p', 'if E*, then believe not-p', etc.

Focusing on these considerations provides considerable support for the Permissivist view. For, plausibly, there are many distinct rational methods of analysing a body of evidence. Take for example the epistemic standards of 'infer to the best explanation' or 'have beliefs that are coherent with your other beliefs'. These are of course distinct epistemic standards; yet, they are both presumably rational. So, if these distinct standards could at points deliver different results given a fixed body of evidence, then this would mean that there can be Permissive cases. Now, is the antecedent of this claim true? As indicated above, the case of disagreement between Palaeontologists gives good reason to think so. For, we can plausibly assume this dispute to arise a result of the parties adopting distinct, yet equally rational methods of analysing the available evidence to draw conclusions. For instance, we might suppose that the different group of Palaeontologists have distinct weighing of certain theoretical virtues—e.g., different ways of balancing

coherence with predictive accuracy. Similarly, we can potentially explain the disagreements between Stalnaker and Lewis on how to flesh out the notion of 'possible worlds' as originating from distinct methods of analysing the evidence as a result of differences on which theoretical virtues are considered more desirable—e.g., with Lewis favouring coherence more than the commonsensicality when compared to Stalnaker. In this sense, we can explain these disagreements as Permissive cases insofar as they can be traced back to distinct, yet equally rational ways of analysing a body of evidence that can render different propositions supported by that evidence.

4.3.2 Permissivism in Philosophy

So far, I have provided motivations for thinking that there could be Permissive cases of disagreement. In effect, these considerations suggest that parties to a dispute could be rational in maintaining their respective diverging positions. Now, what's important to note here is that this raises the question of whether the disagreements uncovered by evidence of cognitive diversity could also be thus permissive. If so, this would provide a direct rebuttal of the first premise of the CICD—i.e., the claim that findings of cognitive diversity about a thought-experiment T provide strong undefeated evidence that intuitive judgements about T are irrational. I can think of two ways Critics could respond to this proposal:

1. Deny Permissivism outright—i.e., to defend the position commonly known as 'Uniqueness'; *or*
2. Accept that there are Permissive cases, but contend that the cases of demographic variation in intuitive judgements about thought-experiments do not admit of more than one rational doxastic attitude.

The first option gives a straightforward way to undercut the challenge raised against the CICD. Denying that there could be any such Permissive cases obviously precludes the possibility that any disputes about what obtains in philosophical cases could allow for more than one rational doxastic attitude. However, there are very good reasons to resist this option. The chief reason is that there is a great deal of controversy about which of Uniqueness or Permissivism is true—each proposal having its own long list of proponents (for a review: Kopec and Titelbaum, 2016). So, if Critics wish to articulate the CICD by way of a defence of Uniqueness, then this challenge significantly loses its bite. After all, taking this route would leave open a rather simple response to this challenge: viz., appeal to the considerations that are often taken to motivate Permissive cases.

The second option is far more promising. For, even if one leaves it open that there could be permissive cases of peer disagreement (such as that between the Palaeontologists, or that between Stalnaker and Lewis), this need not entail that *every* case of peer disagreement admits of more than one rational doxastic attitude. Indeed, there are quite good independent reasons to think that such an extremely permissive view is ultimately untenable (Christensen, 2016). In this sense, Critics could seek to establish that even if the case of the Palaeontologist or of the philosophers on the nature of possible worlds is indeed Permissive, they are importantly distinct from the disputes cutting across demographic lines. I take it that the most promising approach here would be to explain that, although these disputes are explained in terms of distinct yet rational approaches to evaluation of evidence, the case of disagreement uncovered by findings of cognitive diversity seems to lack this feature. After all, in these cases, the only identifiable difference

between the parties of the dispute is their *demography*; and, if this is agreed to be a problematic irrelevant influence on belief, then there does not appear to be much to indicate that the disagreement could ever be rational.

However, I contend that even if we can trace disagreements between peers back to influences of arbitrary demographic factors, this is not necessarily indicative of irrationality on the part of any of those involved in the dispute. Rather, I contend that some such cases of disagreement are quite similar to the Permissive cases of peer disagreement detailed above (that is, the case of Palaeontologists who differ with respect to their views on what led to the death of the dinosaurs, or of the philosophical disputes between expert philosophers). Specifically, as in these cases, evidence of disagreement about philosophical cases along demographic lines can indeed be explained as the result of distinct *rational* methods of analysing the evidence—thus making them Permissive cases.

As a first step to developing this argument, it will be useful to briefly motivate the idea that some epistemic norms might be beholden to local conditions. Brown's (2013) comments on this topic are here instructive. She suggests that the epistemic norms governing belief and practical reasoning (i.e., norms of the form 'if E, believe p') might be primarily a matter of social coordination: i.e., they are norms that are defined by the community in order to respond to their common concerns and limitations. To illustrate, Brown considers the challenge posed by the *threshold problem* for knowledge—i.e., the question of what determines the required strength of a subject's epistemic position to ascribed knowledge. Now, it is plausible to think of the threshold problem as an issue that requires a trade-off between value of truth against the dis-value of falsehood (Foley, 1992). That is, settling the precise conditions for knowledge will require making a choice in weighing the badness of acquiring bad beliefs in opposition to the relative advantage of coming to hold true beliefs. Building on such considerations, Brown (2013) entertains the hypothesis that two communities could come to distinct solutions to the threshold problem as a function of just how they strike this balance. For instance, a community that is very risk-averse—where the cost of acquiring false beliefs is perceived as very bad—will probably be more prudent in its knowledge-attributions. By contrast, communities where the risk of error is weighed less strongly might endorse a lower threshold for knowledge, and will tend to attribute knowledge more easily than the previous group.

In effect, the above example suggests that these two communities will have distinct norms for belief and practical reasoning. Thus, when given a same body of evidence, their respective ways of analysing the evidence to draw conclusions from it will be distinct. In particular, we can expect one group's norms to lead them to attribute knowledge more often than the other. Now, assuming members of these two communities may indeed hold such distinct norms, then it is sensible to expect that they might diverge in some of their intuitive judgements about philosophical thought-experiments. Moreover, assuming such divergences in intuitive judgements do arise, this will plausibly be attributed to influence of arbitrary demographic factors—viz., the particular community a member belongs to. After all, a given member of one such community will have the intuitions she does *because* she belongs to that specific community, which lead her to endorse the norms she does.

However, I submit that even if we assume that communities may disagree in this way, this is arguably

consistent with the claim that they are both rational in their respective positions. For, if norms governing belief and practical reasoning are indeed a matter of social coordination, then it is possible that each group's distinct norms could turn out to be rational. In particular, we can imagine that each group's weighing of the trade-off between the badness of acquiring bad beliefs in opposition to the relative advantage of coming to hold true beliefs adequately attends to the demands and concerns of that particular group. If so, then their respective norms are indeed rational. After all, these norms will help members of these distinct communities to fulfil their respective goals and objectives. As such, I submit that the above divergence in the intuitive judgements of these two communities could plausibly be taken to be a Permissive case of disagreement.

At this point, it is important to clarify the dialectical purpose of the discussion so far. First, it should again be emphasised that I am not defending that *every* philosophical case will be Permissive in the way described above (such that any intuitive judgement is as rational as the next). Rather, my aim in presenting this example is to detail at least one such case where evidence of cognitive diversity would not *uncontroversially* show that any one of the disagreeing parties is thereby irrational. In effect, these considerations echo the suggestion from the previous section that influences of arbitrary factors need not constitute problematic influences on judgement. Recall, we motivated this idea by way of an example of how having insomnia might be considered an irrelevant (yet untroubling) factor in my coming to believe that the corner shop is open at 3 am. And, as explained above, the central reason for this is that facts about my insomnia in no way speak to the irrationality of my belief about the corner shop's opening hours. Similarly, as the above example illustrates, the fact that one's culture caused one to have a method of reasoning that differs from those of another group's doesn't speak to the irrationality of those judgements. In this way, I submit that there are reasons to resist the first premise of the CICD—i.e., the claim that evidence of cognitive diversity about a thought-experiment T provides strong reasons to think that intuitive judgements about T are irrational.[14]

4.3.3 Multiple Concepts?

So far, I have argued that if there are Permissive cases of cognitive diversity, then there are reasons to doubt the CICD. I then described an example of how two communities might have reasonable disagreements in their intuitive judgements about knowledge. So, I concluded that the first premise of the CICD is questionable. Now, one plausible suggestion here is that these two communities do not genuinely disagree; rather, they employ *distinct concepts of knowledge*.[15] In this sense, my arguments against CICD boil down to a familiar line of reply to the experimentalist challenge: viz, that evidence of cognitive diversity in philosophy can be explained away as cases where people are simply talking past each other (see, e.g., Sosa, 2009). Although seemingly plausible, I will argue that this is not the case. My main contention is

[14] One important question at this point regards the matter of who has the burden of proof here. For instance, the Critic could simply claim that it is on the Defender to show that the cases in which they find evidence of demographic variation are instances of such Permissive cases. I will not say much about this question here. My aim in this chapter is merely to show that it is at least not straightforward that if there is evidence of demographic variation, then that intuition is thereby unsuitable for use in philosophical inquiry.

[15] For a formulation of this proposal, see Sosa (2007a)

that there are good reasons to think that the disagreement detailed above does indeed concern uses of a *single* concept of knowledge. To buttress this claim, I now briefly discuss one set of arguments to this effect. Before proceeding, I should make it clear that the aim of this discussion is not to provide a detailed and thorough defence of such arguments; rather, my purpose is only to detail why it is reasonable to think that the two communities described above can indeed be said to share a same concept of knowledge. Building on these considerations, I propose that an argument to the effect that the disagreement is indeed about *distinct* concepts would be controversial.

As a first step to developing this argument, we must first settle on an approach to defining the concept of knowledge. By all accounts, this is no easy task: many recent debates in epistemology can be aptly characterised as turning on precisely this issue (for a discussion: Ichikawa and Steup, 2018). However, one set of recent proposals promises to make headway insofar as it aims to find a middle ground amongst such many dissenting views. The unifying thread of all such proposals is the idea that we can better define the contours of our concept of knowledge by clarifying the needs and demands this concept satisfies. Otherwise put, this approach seeks to define the concept of knowledge by determining what is its *function* (i.e., it's purpose or point) in the epistemic lives of those using that concept (for a collection,: Henderson and Greco, 2015).

Taking a functional analysis of the concept of knowledge allows for an explanation of how two communities can be said to share a single concept of knowledge, yet still diverge in how they employ that concept. To begin explaining, it is important to clarify that on such functional analyses, two communities will be said to share a concept of knowledge if their long-standing patterns of use of that concept turn out to be *functionally* similar. That is, functional analyses will yield the verdict that communities have a common concept of knowledge if that concept plays similar roles in the epistemic lives of their members. Now, we can stipulate this to be the case in the example detailed in the previous section. For instance, following the currently most well-developed and influential functional analyses of the concept of knowledge, let's determine that the members of these communities use this concept with the function of identifying *good informants* (see, e.g.,: Craig, 1990; Henderson and Greco, 2015; Kusch and McKenna, 2018). Thus, if we analyse the patterns of use of the concept of knowledge in these communities, we will find that their epistemic evaluations using that concept will function to identify reliable informants.

However, note that this functional analysis leaves it open that two communities may indeed share a concept of knowledge yet hold slightly distinct standards for determining what counts as a reliable sources of information. That is, there might be differences in *how* epistemically reliable an agent will need to be in order to be regarded as a 'knower' in these distinct communities.[16] One way to flesh out this point is to consider the example detailed in the previous section. Recall, there we detailed the case of two communities that come to differing thresholds for knowledge as a result of their distinct solutions to the threshold problem: the more risk-averse community endorsed a higher threshold for knowledge, whereas the less risk-averse community endorsed a lower threshold. Building on this proposal, it is reasonable

[16] Notably, these views can be taken to motivate Contextualism about knowledge (see, e.g., Henderson, 2009; Henderson, 2011; Hannon, 2013); however, they certainly do not entail this view (see Hannon, 2017, fn. 23).

to expect that even if these communities both use knowledge to pick out reliable informants, they might still disagree in their knowledge attributions. Whereas the more risk-averse community will be more cautious in their knowledge-attributions, the more risk-seeking community will be a bit more liberal in such ascriptions. Thus, we can expect that a subject S might be said to know that *p* in one community, yet be said to not know that *p* in another. But, note that even if we stipulate that these communities diverge in this way, this is still compatible with the claim that these two communities *generally* agree about what counts as a reliable source of information. Otherwise put, we can say that the differences between these communities might be very small—indeed, only marginal. However, even such minute differences could pull their evaluations of knowledge attributions apart in very tricky and difficult cases. And critically, note that the kinds of cases that play a central role in the epistemological literature are precisely of such a sort: they are typically unclear and boundary cases that work 'at the margins' of our concepts. Thus, even if we can expect there to be significant overlap between how these communities employ the concept of knowledge, the minute differences in their epistemic standards can be revealed when they evaluate such borderline cases.

At this point, it is useful to pause to again clarify the dialectical purpose of the considerations in this section. Notably, my aim is not to provide a thorough defence of functional analyses of knowledge. Instead, the purpose of these considerations is simply to explain why it is reasonable to think that two communities can indeed share a concept of knowledge, yet disagree in their evaluations of knowledge attributions. This then suggests that the above putative cases of Permissive cognitive diversity is not simply a case where parties to the dispute simply have distinct concepts and are thereby talking past each other.

4.4 Against the CFCD

In this section, I argue against the argument I have denominated the Challenge of Falsity from Cognitive Diversity (CFCD). That is:

P1 If there is cognitive diversity about a thought-experiment T, then this provides strong undefeated evidence that intuitive judgements about T are false.

P2 If there is undefeated evidence that a judgement is false, then that judgement does not provide warrant for philosophical positions and claims.

C Therefore, if there is cognitive diversity about a thought-experiment T, then intuitive judgements about T do not provide warrant for philosophical position and claims.

My main contention is that the first premise of this argument fails. To argue for this proposal, I take on a similar strategy to that of the previous section: namely, I will argue that, although seemingly compelling, this premise can be shown to rest on questionable assumptions. Now, it is important to make it clear at the outset that the arguments I develop for this proposal are importantly distinct from those developed in the previous section—which sought to establish that, *pace* Critics, there are cases in which there is more than

one rational intuitive judgement of a philosophical case. For, note that a similar argument for a plurality of *correct* responses to philosophical cases would be much more controversial. Instead, in this section I propose that divergences in the responses of demographic groups about a thought-experiment T is a multifaceted phenomenon: there are many ways in which intuitive judgements can be found to diverge. And, as we will see, in a restricted set of cases there is good reason to think one of the groups is indeed better positioned to judge on the matters under dispute. This effectively suggests that in some cases, even if we find demographic variation in intuitive judgements, this does not provide decisive reasons to think that such intuitive judgements are false.

Typically, philosophers employ thought-experiments in inquiry as *intuition pumps*: that is, they appeal to thought-experiments with views to teasing out a *specific* intuition which they then use to answer a philosophically relevant question. Perhaps the most prominent example of this is Gettier's (1963) use of two thought-experiments to elicit an intuition that a given subject does *not* know a given proposition. Relying on this intuitive judgement, Gettier sought to provide a negative answer to the question 'is knowledge reducible to justified true belief?' Other famous examples are Kripke's (1980) appeal to intuitions about the case of Gödel to answer whether the descriptivist theory of reference was correct, or Jackson's (1982) use of intuitions about the case of Mary to answer whether the mind was reducible to the physical. Let's call this use of thought-experiments as *intuition pumps* as the Standard implementation of thought-experiments.

It is important here for us to appreciate the implications that findings of cognitive diversity have for these standard implementations of thought-experiments. For instance, consider the following scenario: people from culture A have the intuition that the subject in a Gettier case *does* know, whereas people from culture B have the intuition that the subject does *not* know. For the reasons detailed in the previous sections, there does seem to be something troubling about appealing to intuitions about Gettier cases to answer whether knowledge is reducible to justified true belief. After all, even if we assume that people from both culture A and B are rational in their respective intuitive judgements, this stills provides good reason to think that either of those intuitive judgements are false.

So, if we focus exclusively on these standard implementations of thought-experiments, we do seem to have good reason to think that a threat of falsity looms—which, in turn, lends weight to the CFCD. But, what is crucial to point out is that such standard uses are not exhaustive of the ways in which philosophers make use of thought-experiments. To illustrate, consider the following two alternative implementations of thought-experiments (discussed in chapter 3):

Aporetic Uses: Some thought-experiments are used in order to elicit *conflicting judgements* about the same scenario. I will call these Aporetic uses of thought-experiments to indicate that they are meant to give rise to a kind of *puzzlement* that helps to refine inquiry into their subject-matter. A simple example is the case of the ship of Theseus: "is a ship that has had all of its components gradually replaced on the course of a voyage the same ship on arrival as on departure?" Most people feel the pull of both the positive and negative answers: in one sense, it is indeed the same ship; but, in another, it seems that by substituting all of its parts we now have an entirely different ship. This thought-experiment thus

gives rise to a sense of puzzlement about the notion of identity, which can then be put to use in refining inquiry. For instance, if it turns out that the intuitive pull of each answer corresponds to a different concept of identity, then philosophers will have to hone their inquiry accordingly. Newcomb's problem is another instance of an Aporetic use of a thought-experiment. This is a puzzle that elicits two plausible yet inconsistent judgements about rational attitudes (Nozick, 1969). Reflection on this problem has led to animated (and ongoing) discussions about the source of the puzzlement; it has promoted philosophers to draw novel distinctions and formulate questions that attempt make inquiry into rational attitudes more precise. (For instance, Newcomb's paradox shows that there are cases where the decision-theoretic principle of dominance gives advice that conflicts with the principle of expected utility-maximization; this is a surprising result as it dissociates our judgements about rationality from the predictions of standard decision-theory.) In sum, Aporetic uses of thought-experiments significantly improve philosophical inquiry, to the extent that they tease out important distinctions that help us to better understand central philosophical notions.

Critical Uses: Some thought-experiments seek to elicit intuitions in order to scrutinise which aspects of the specified case elicit them. Insights into which features of a case are responsible for eliciting certain intuitions can then be used in assessing the evidential value of those intuitions. One example of such Critical uses of thought-experiments is Gendler's (2010) close scrutiny of a range of scenarios used to investigate the concept of personal identity. From this detailed analysis, she concludes that some of the most prominent thought-experiments used in the literature are just too far divorced from the situations in which we ordinarily apply the concept of personal identity. Her central contention is that our intuitions about such cases are thereby unlikely to track features essential to the concept of personal identity, making them useless for philosophical theorising. Other cases of Critical use aim to not only show an intuition to be unwarranted, but to argue that another (opposing) intuitive judgement is correct. One example of such a Critical use is Jackson's (2003) *later* appeal to the case of Mary the Neuroscientist, where he attempts to diagnose *why* this case tends to elicit the intuition that Mary does indeed learn something new when she sees the red rose for the first time. Jackson invokes this intuition as a foil for his argument to the effect that this judgement turns on a mistaken conception of the nature of perceptual experience—viz., one which erroneously characterises experience in terms of *what it's like*. Building on this argument, he proposes that we can come to see why the right intuition to have in the case of Mary the Neuroscientist is that Mary does not learn something new.

Prising apart Aporetic and Critical uses underscores that the focus on standard uses of thought-experiments fails to capture other important aspects of our practices. And, more importantly, focusing on this broader set of practices helps us to appreciate that demographic variation in intuitive judgements about a thought-experiment T need not be taken to indicate that *all* intuitive judgements about T are equally likely to be incorrect. To illustrate, let us focus on the thought-experiment of the Ship of Theseus. Suppose that empirical findings reveal a significant degree of demographic variation in intuitions about this thought experiment. For the sake of illustration, suppose such findings reveal that all people from culture A have an intuitive judgement that yes, the ship that returns to port is indeed the Ship of Theseus, whereas people from culture B have an intuitive judgement that it is not. Now, notice that this wouldn't necessarily be

indicative that either of these intuitive judgements is thereby likely to be inaccurate. After all, in the Aporetic use of this thought-experiment we would *expect* conflicts in intuitive judgements to arise. Indeed, the use of this case is generally to pull apart these distinct intuitions. Now, if the conflict that is expected to arise track demographic factors, and assuming that no demographic group is in better epistemic standing than the other, then the finding of cognitive diversity would provide no reason to be suspicious of any of the intuitive judgements. At most, this would provide an interesting piece of data that informs further inquiry, which is precisely the aim of such Aporetic uses of thought-experiments. Thus, I suggest that it is not the case that cognitive diversity about thought-experiment T thereby indicates that intuitive judgements about T are likely to be inaccurate.

Focusing on Critical uses of thought-experiments helps to buttress this proposal. Identifying Critical implementations of thought-experiments underscores that the focus on standard uses of thought-experiments fails to capture other important aspects of our practices. And, more importantly, focusing on this broader set of practices helps us to appreciate that demographic variation in intuitive judgements about a thought-experiment T need not be taken to indicate that *all* intuitive judgements about T are equally likely to be false. The reason for this is that many Critical uses of thought-experiments underscore divergences in intuitive judgements about a thought-experiment. But, importantly, in many such Critical uses, the aim is to explain *just why* one such intuitive judgement is in better epistemic standing than another. And, if such a Critical use is successful in doing so, even if the relevant divergence in intuitive judgements about a given thought-experiment tracks demographic factors, this would only give us reason to think that one such demographic group is in better epistemic standing to judge on these relevant matters. To flesh out these claims, I now briefly describe one recent perspicuous Critical use of thought-experiments in Epistemology.

Epistemic Contextualism is (roughly) the view that knowledge ascriptions are sensitive to considerations about the stakes, as well as to whether possibilities of error are made salient in the context of the attribution (DeRose, 1992; Cohen, 1999; Fantl and McGrath, 2002). One of the central lines of support for Contextualism about knowledge is that people often attribute knowledge to a subject if no possibilities of error are made salient in that context, but will often retract such attributions once such possibilities are brought up. To illustrate, consider the widely discussed 'bank case' in which Jane and her wife are debating whether to go to the bank to deposit their cheques, or wait till the next day—a Saturday—to do so. Jane claims she knows that the bank will be open on Saturday because it was open on Saturday two weeks ago. Question: does Jane know that the bank is open on Saturday? Many people answer this question in the affirmative. However, imagine that Jane's wife then reminds her that banks sometimes change their hours, and that if she doesn't deposit the cheque before Monday they might lose the mortgage of their house. At this point, people often deny that Jane might actually know that the bank is open on Saturday. This pattern of shift in people's evaluation have been found in a number of empirical studies, and many have thus claimed that the folk have such 'contextualist intuitions' about knowledge-attributions (Sripada and Stanley, 2012; Pinillos, 2012; Pinillos and Simpson, 2014; cf. Rose, Machery, and al, forthcoming).

Invariantists deny that knowledge is sensitive to stakes or to possibilities of error. Crucial for current

purposes, one recently prominent strategy to develop this conclusion proceeds by way of a Critical use of the thought-experiments that are taken to motivate contextualism. In more detail, these arguments seek to develop a pragmatic explanation of contextualist intuitions, whereby they are taken to reflect the conversational *propriety* of knowledge attributions in those cases—but not the truth of whether the subject knows or does not know in that case. Brown (2006) provides what is perhaps one of the most detailed defences of this position.[17] She argues that contextualist intuitions are due to people respecting a conversational rule of Relevance. To explain, she make recourse to one of Grice's examples of how this rule operates in conversation:

Suppose that a man is standing beside a car in obvious need of repair and asks a passer-by, 'Is there a garage nearby?' In this situation what's conversationally relevant is not merely whether there is a garage nearby, but whether it's open. As a result Grice argues that although the response 'Yes, there is one round the corner' literally states only that there is a garage there, it would pragmatically convey that the garage is open. For, if the passer-by did not think that it is possible that the garage is open then her utterance would infringe the maxim 'Be Relevant'. Thus, the passer-by's utterance would be intuitively incorrect if she knows that the garage is shut. Rather, in such a scenario, and assuming that she believes that there is no other nearby open garage, it would be correct for her to reply 'No, there's no garage nearby'. While this utterance is literally false, it pragmatically conveys the true claim that there is no open garage nearby.

Brown (ibid.) argues that this example of how the Relevance rule can make false statements seem appropriate in certain conversational contexts is suggestive for understanding contextualist intuitions in the bank case (discussed above). Specifically, she suggests that as it is very important in the high stakes case that the cheque be deposited before Monday, then what is *relevant* for the conversation is that one be in an *extremely* strong epistemic position with respect to the bank's opening times—not *merely* whether one knows that the bank is open on Saturday. Thus, attributions of knowledge in these contexts *pragmatically* convey the information that the subject of the bank case is in such a strong epistemic position. But, independently of whether the subject in the scenario does actually know the bank's opening times, attributing knowledge to her in this case would appear to violate the rule of Relevance, as it is not really clear that the subject in that case is really in such a strong epistemic position.

We need not belabour too much on the details of this argument. Rather, there are three points that are here worth noting about this Critical use of thought-experiments. The first is that the above Critical use of thought-experiments about bank cases arguably highlights a case of cognitive diversity with respect to these thought-experiments. For, as stated above, empirical findings confirm that *non-philosophers* generally tend to have contextualist intuitions in response to thought-experiments like the bank case detailed above. In this sense, it is natural to expect that the great majority of people *without philosophical training* will issue a contextualist intuition when presented with these cases.[18] Furthermore, it is reasonable to think that Brown's opposing intuitive judgement about the bank cases is in great part informed by previous philosophical training. Of course, it is entirely possible that someone without any such training

[17] For a similar strategy to argue for Invariantism, see Rysiew (2001).
[18] Cf. Rose et al. (forthcoming).

could share these intuitive judgements. But, what is important to note here is that Brown's reasons for maintaining that intuitive judgement are most plausibly due to a good understanding of the philosophical literature on the pragmatics/semantics distinction, and ability to marshal these considerations to the debates about Contextualism. As such, we can then naturally trace back the divergence in these intuitive judgements to features of demography: viz., the degree of philosophical training.

Second, I propose that the above Critical use of the bank cases provides good reason to think that one of these intuitive judgements is actually in better epistemic standing than the other. For, as the above discussion of the rule of Relevance demonstrates, contextualist intuitions are likely to be (misleadingly) influenced by pragmatic considerations. In particular, these considerations show that, although accurately reflecting wide-spread conversational rules, contextualist intuitions are nevertheless false. In this sense, there is good indication that non-contextualist intuitive judgements about the bank cases are in better epistemic standing. Of course, this is not to say that these are decisive reasons for thinking as much. But, given the plausibility of the pragmatic explanation of contextualist intuitions, then the conflict in intuitive judgements highlighted by the above use of this Critical use of bank cases does not motivate the conclusion that intuitive judgements about this case are equally likely to be false.

Lastly, the considerations in this section show that even if there is cognitive diversity with respect to a thought-experiment that is *typically* used in a Standard implementation (e.g., Gettier or Bank cases), this does not show that it is problematic to appeal to intuitive judgements about that thought-experiment in philosophy. For, as the above Critical use of the bank case demonstrates, considerations about what the details of this case pragmatically convey reveals that one of those diverging intuitions might be in better epistemic standing. And note that a similar conclusion is suggested by a recent analysis Max Deutsch (2009) gives of a study that found cultural variation in intuitions about a modified version of Kripke's famous Gödel case (Machery et al., 2004). Deutsch argues that in the case used in this study, the use of the sentence "When John uses the name "Gödel", he is talking about the man who stole the proof" pragmatically conveys that John *intends* to refer to the man who stole the proof when using "Gödel". Deutsch's analysis can effectively be taken to show that the findings of cognitive diversity on intuitions about this case are possibly due to people from certain cultures being more likely to misled by this pragmatic implication than others. And, note that this analysis could be taken to show that one set of intuitions is in better epistemic standing—such that it would be permissible to rely on those intuitions in philosophical inquiry.

The considerations in this section thus indicate that influences of arbitrary factors need not constitute problematic influences on judgement. We first motivated this idea via an example of how my insomnia might be considered an irrelevant, yet untroubling factor in my belief about the corner shop's opening hours (see sec. 4.2.1). Again, the central reason for this is that facts about my insomnia in no way speak to the falsity of my belief about the corner shop's opening hours. Similarly, as the above discussion illustrates, the fact that one's intuitive judgements are caused by demographic factors (viz., training in philosophy) doesn't speak to the falsity of that judgement. In this way, I submit that there are reasons to resist the first premise of the CFCD—i.e., the claim that evidence of cognitive diversity about a thought-experiment T

provides strong reasons to think that intuitive judgements about T are false.

4.5 Conclusion

I have been evaluating the viability of the Challenge from Cognitive Diversity (CCD). Through a detailed analysis of the considerations and epistemological principles driving this challenge, I have proposed that there are reasons to doubt that it succeeds outright. In order to articulate this claim, I first appealed to recent epistemological disputes about the phenomenon of contingency anxiety. From these debates, I drew a distinction between two versions of the CCD: the first focusing on the threat of irrationality arising from evidence of cognitive diversity (CICD) and the second on the threat of falsity (CFCD). I then argued that both these formulations fail as they require defence of controversial positions. In particular, the CICD was shown to require a defence of the controversial view known as 'Uniqueness'—the idea that there is a single rational doxastic attitude to take towards a body of evidence. The CFCD, on the other hand, requires defending a very narrow view of how thought-experiments are used in philosophical inquiry. The central upshot of these arguments is that the CCD is not as compelling as Critics would have us believe. As such, I contend that even if Critics amass significant evidence of cross-cultural variability in philosophical intuitions, this is still insufficient to buttress a radical version of the experimentalist challenge.

Chapter Five

Intuitions and Contrastivism about Interrogative Understanding

The arguments from previous chapters suggest that intuitions can provide warrant for philosophical positions and claims; but, that philosophers should engage with empirical findings from cognitive psychology to determine when to trust their intuitions, and when to refrain from appealing to them in inquiry. However, the arguments I developed in those chapters did not explain the kind of epistemic standing intuitions afford with respect to their contents. In this chapter, I address this issue. My aim is to *explore* the idea that veridical intuitions give rise to *interrogative* understanding—such that a veridical intuition that p will provide one with understanding of why p. To investigate this proposal, I take on a two-step approach. First, I argue for a novel *contrastive* view of interrogative understanding—according to which to understand why p is to understand why p *rather than* some alternative, q. I then show that focusing on the contrastive nature of interrogative understanding helps to explain why intuitions are poised to provide warrant for philosophical positions and claims.

Here is a more detailed plan of the chapter. I first provide a quick overview of the central features of interrogative understanding and motivate the idea that intuitions *can* give rise to interrogative understanding of their contents (sec. 5.1). This will underscore the relevance of investigations into the nature of interrogative understanding for the study of the epistemology of intuitions. I then argue for the novel view according to which interrogative understanding is *contrastive* (sec. 5.2), and defend it from objections (sec. 5.3). Lastly, I propose that focus on the contrastive nature of interrogative understanding proves useful to flesh out the claim that intuitions provide epistemic warrant for philosophical positions and claims (sec. 5.4).

5.1 Interrogative Understanding and Intuitions

5.1.1 Interrogative understanding

The philosophical study of *understanding* distinguishes between the following ascriptions of this epistemic state:

- Arjun understands Spanish. (Linguistic understanding)
- Muhammad understands Physics. (Objectual Understanding)
- Raissa understands that 'America' refers to North America. (Propositional Understanding)
- Claire understands why the Earth revolves around the Sun. (Interrogative Understanding)

In this chapter, I will focus on the latter (interrogative) kind of understanding. Inquiry into the nature of interrogative understanding focuses predominantly on our patterns of ascriptions of this epistemic state (both to others and to ourselves). Typically, these investigations proceed first by examining a number of (hypothetical and real) scenarios and examining if a subject S in those circumstances can plausibly be said to have interrogative understanding of a given fact p. The central motivation for this approach is the assumption that our evaluations of 'S understands why p' claims shed light on what conditions S must satisfy to have interrogative understanding about p. Thus, on this approach, if we wish to gain explanatory purchase on what is required to understand why crows fly, then we should attend to the circumstances under which we attribute (or don't attribute) understanding of why crows fly to others or to ourselves. Attempts to define the nature of interrogative understanding through this approach have proven notoriously divisive (see, e.g., Grimm, 2017).[1] For the most part, these discussions focus on the relation of interrogative understanding with propositional knowledge, epistemic justification, and luck, and on whether understanding why p requires abilities to explain why p. In this chapter, I will remain neutral on *all* such disputes. Instead, I focus on features of this epistemic state that are common amongst the main views on the nature of interrogative understanding. I detail these below.

Attributions of the form 'S understands why p' are commonly regarded to ascribe a particular epistemic relation between S and a *correct* answer to a why-question (i.e., questions of the sort "why p?").[2] In effect, this means that interrogative understanding is factive. Thus, one cannot understand why air is fluid if this is false. Moreover, to understand why air is a fluid, one must have a *correct* answer to the question 'why is air a fluid?'—and not just *any* answer. Similarly, to understand why Venezuela is in crisis, one needs to possess a *correct* answer to the question 'why is Venezuela in a crisis?'—not a faulty account of why this is so.

Now, it is quite clear that *merely* possessing the answer to the question 'why p?' does not suffice for understanding why p. To illustrate, consider the case where I report the right answer to the question 'why do crows have feathers?', by stating the relevant evolutionary facts. But, suppose that because of my

[1] For some recent work on interrogative understanding, see Hills, (2015) and Sliwa (Sliwa, 2015; Sliwa, 2017).

[2] Some might wonder whether this commits one to a propositional view of the object of interrogative understanding. I contend they do not, as I leave it open that one can reasonably endorse a non-propositional view of the semantics of why-questions (see, e.g., Ginzburg, 1995a; Ginzburg, 1995b).

firm religious beliefs I deny this evolutionary explanation. It is quite natural to think that I thereby fail to understand why crows have feathers. This observation suggests that if S understands why *p*, then S must both have an answer to why *p* and an appropriate doxastic attitude (e.g., a belief or judgement) towards that answer.[3]

Understanding of why *p* is, however, not just a mere doxastic attitude towards a correct answer to the question 'why *p*?'. After all, one may genuinely believe an explanation for why *p* yet fail to understand why *p*. For instance, consider the case where I come to believe the explanation for why electrons spin in opposite directions by route memorisation. Suppose I read off from Wikipedia that this is because the Pauli exclusion principle rules out that these particles can have the same quantum numbers, and therefore the only way for electrons to occupy the same orbital is for them to differ in the spin quantum number. Now, imagine that because I (correctly) judge Wikipedia to be a reliable source, I then come to have a justified belief (and knowledge) of why electrons spin in opposite directions. Now, it is clear that simply believing this explanation in this way may fall short of full understanding of why electrons have opposite spins. Thus, it is commonly agreed that interrogative understanding requires—but is not reducible to—a doxastic attitude towards the right answer to a why-question.

There is an ongoing heated debate about which conditions—besides having a doxastic attitude towards a correct answer to the relevant why-question—are required for interrogative understanding. Nevertheless, parties to this disagreement accept that understanding why *p* involves a distinctive (broadly defined) 'grasp' of a body of information that answers the question 'why *p*?'.[4] Unfortunately, it is far from clear what precisely this mental state of 'grasping' amounts to. To clarify the relevant notion of 'grasp' at stake in debates about interrogative understanding, many have appealed to an analogy with the grasping of objects. For instance, Hills suggests that:

> The best way of thinking of it is by an analogy with grasping a ball or cup of tea or similar. If you grasp a ball, you have it under your control. You can manipulate it, move it, turn it round, and so on, that is you (normally) have a set of practical abilities or practical know how, which you can exercise if you choose.
>
> (Hills, 2015, p. 663)

Hills builds on this proposal to elucidate her claim that understanding why *p* "requires a grasp of the reason why *p*, or more precisely, a grasp of the relationship between *p* and q" (ibid., p. 663). In more detail, she contends that understanding why *p* involves a grasp of the relationship between *p* and *q*, such that one has that relationship under 'one's control'. That is:

[3] For sake of simplicity, I will henceforth speak of interrogative understanding as *strictly* related to beliefs; however, I will remain neutral on the question of whether interrogative understanding *entails* a belief. For, plausibly, there are doxastic states that are unlike beliefs and judgements (e.g., acceptances, rational endorsements), and which can also give rise to interrogative understanding.

[4] Although not all parties to this dispute employ the term 'grasp' in quite the same way. For instance, some speak of a grasp of propositions (e.g., Hills, 2015), whereas others reserve grasp to non-propositional contents (e.g., Grimm, 2014).

> You have a set of abilities or know-how relevant to [this relationship], which you can exercise if you choose. For instance, if you understand why *p*, you can give an explanation of why *p* and you can do the same in similar cases. If you find out that *q* (where *q* is why *p*), you can draw the conclusion that *p* (or that probably *p*, if *q* only makes *p* probable). And you can do the same in similar cases.
>
> (Hills, 2015, p. 663)

Grimm appeals to a similar analogy of grasping objects to account for the nature of interrogative understanding. He suggests that the distinctive mental state of 'grasp' involved in this epistemic state is similar to that of "grasping a thing when we are able to manipulate or tinker with a thing", such as when we grasp "a simple lever system by manually switching the lever from one position to another" (Grimm, 2011, p. 89). Building on this suggestion, Grimm proposes that the *mental* grasp of such a system would "bring into play something like a modal sense or ability," whereby one registers not just how things are in that system, but how "changes in one part of the system will lead (or fail to lead) to changes in another part" (ibid., p. 89). In effect, Grimm's proposal amounts to the claim that grasping a fact *p* involves a type of *awareness* of the many causal factors that come together to bring about *p*. Thus, to understand why crows fly is to be aware of, say, the different evolutionary facts that make it the case that crows fly, and to recognise how changes in those facts would affect whether or not crows fly.

These tentative characterisations of the mental state of grasping are not offered here as being exhaustive or satisfactory. Indeed, for current purposes I will remain entirely neutral on the issue of what is the best account of grasping. Rather, my aim in discussing these proposals is that they help to articulate the view that I seek to explore in this chapter: namely, the idea that intuitions can enable interrogative understanding of their contents. With these characterisations in mind, I now turn to an articulation of this view.

5.1.2 Intuitions and interrogative understanding

In this section, I motivate the idea that intuitions can enable interrogative understanding of their contents. According to this view, a *veridical* intuition that *p* affords understanding of *why p*. Notably, my aim in this section is not to provide a thorough *defence* of this view. Rather, I seek only to detail how many prominent accounts of the nature of intuition suggest this view. In this way, I take the discussions in this section to reveal that inquiry into the nature of interrogative understanding is germane to debates about the epistemology of intuitions.

The proposal that there is an intrinsic connection between intuition and understanding is in no way novel. Indeed, early writings on the nature of intuition by Leibniz and Descartes suggest this link. For instance, Leibniz claims that "[when] my mind understands at once and distinctly all the primitive ingredients of a conception, then we have intuitive knowledge" (Leibniz, 1973/1686, p. 42). Leibniz's proposal is that intuition involves more than just a belief, or an appreciation that a proposition is true. Rather, it involves an apprehension of the concepts involved in that proposition, such that one 'perceives' or 'grasps' (in

the factive sense) the identity of those concepts. In effect, Leibniz's account of intuition is that intuitive knowledge involves a direct 'grasp' of its contents in a sense that is analogous to the grasp of material *objects* (familiar from recent accounts of interrogative understanding).[5] Descartes suggests a similar link between intuition and understanding when he characterises intuition as a state in which the mind "sees, feels and handles" a proposition in a way that makes it clear why it is true (Descartes, 1991, p. 331). Descartes effectively suggests that intuition allows one to directly grasp or see the 'portions of reality' (i.e., the abstract objects) that make this proposition true. Again, the relevant notion of 'grasp' of ideas which Descartes appeals to in defining intuitions is that of grasping objects, similar to that used in recent definitions of the nature of interrogative understanding. Thus, like Leibniz, Descartes also posits a link between intuition and the mental state of grasping that is characteristic of interrogative understanding.

Beyond such historical examples, some influential contemporary accounts of intuition also posit a link between intuition and interrogative understanding. These contemporary accounts also characterise intuitions in terms of a mental state of grasping that is strikingly similar to that used to define the state of interrogative understanding. Now, it is noteworthy that these recent accounts are importantly distinct from those of Leibniz and Descartes, which define intuitions as success-entailing (i.e., as factive). Contemporary accounts, by contrast, take it that an intuition that p can be false. In this sense, they suggest a weaker view according to which *veridical* intuitions enable interrogative understanding of their contents—whereas non-veridical intuitions at most *appear* to do so. For instance, Bonjour suggests that intuitions "at least purport to reveal not just *that* the claim is or must be true but also, at some level, *why* this is and must be so" (BonJour, 2013, p. 99).[6] To illustrate, Bonjour considers the proposition that '2+3=5'. His proposal is that anyone who thinks carefully about this proposition

> will be able to see or grasp immediately that it must be true, that it is true in any possible world or situation – and that the same thing is also true of indefinitely many further examples of these sorts and others.
>
> (ibid., p. 99)

On a similar note, Chudnoff claims that to have an intuition that p is not just to have a doxastic attitude towards p; rather, it is to have a mental state that (at least) seems to make one aware of why p is true.[7] For instance, Chudnoff suggests the following illustration:

> I consider a few cases and then (C) [the claim that all natural numbers have a finite number of predecessors] seems true to me. [...] When my experience represents that (C) is true it does not just make it seem to me that (C) is true. In having that experience I also seem to be aware of why (C) is true. The structure of the natural numbers seems somehow present to mind

[5] For an enlightening discussion of Leibniz's view on intuition, see Picon, (2008).
[6] Italics are mine.
[7] For contrastivism about causation and freedom, see Blaauw (2012); regarding reasons, see Snedegar (2017); and for contrastivism about knowledge, see Schaffer (2004).

(Chudnoff, 2013, p. 62)

Chudnoff argues further that if such an intuition is veridical, then it will indeed make one aware of the abstract objects that reveal why (C) is true (rather than just make it seem that one is so aware). In this sense, he suggests that a veridical intuition allows one to "see" that which makes it true.

And lastly, Bengson also posits a link between intuition and interrogative understanding.[8] To motivate this proposal, Bengson considers the following case:

> *The Cogito*. The meditator takes herself to know many things, including that she exists and that she thinks. Still, there is a sense in which at the outset of her meditations the meditator does not understand these things. At some point, however, a shift occurs. She considers the evil genius scenario and *sees* that, whatever else she might be wrong about, it is impossible for her thought that she exists to be false.
>
> (Bengson, 2015a, p. 646)

Bengson explains that this 'shift' in the relevant subject in this vignette as the result of having a veridical

> modal intuition that, whatever else she might be wrong about, it is impossible for her thought that she exists to be false. As a result her conception changes, the elements coalescing in a way that they did not previously, as her conception now identifies a pertinent substantive connection—a kind of guarantee—between her thinking and her existing that was previously unidentified. In this way, she puts the pieces together and comes to conceive of herself as being such that her thinking guarantees her existing. [...] In this case, the epistemic shift, a qualitative *improvement in the meditator's understanding*, occurs in virtue of an intuition [...][9]
>
> (ibid., p. 657)

Common to all these contemporary proposals is the idea that a *veridical* intuition allows one to 'see' or 'grasp' why its contents are true. And critically, the relevant sense of 'grasp' or 'see' here is strikingly similar to that which Grimm describes in understanding why p, whereby one has a 'modal sense or ability' to register not just how things are, but how certain facts come together to bring about p. It is in this sense that Bonjour claims that an intuition that 2+3=5 enables one to 'grasp immediately that it *must* be true, and that it is true in *any possible world*'. Chudnoff also appeals to this notion of a modal ability to

[8] Bengson does express reservation about establishing a strong link between intuition and *some forms* of interrogative understanding. But, the account of interrogative understanding he considers is a narrow one, wherein understanding why p is tightly linked to possessing an explanation for why p. However, the epistemic state of interrogative understanding I am considering here is not linked to explanation in this way—and can thus be aptly characterised as what Bengson denominates an 'innocent' (non-explanatory) intuition.

[9] Italics are mine.

'register how elements in a system behave' when he claims that his intuition that (C) makes him aware of the structure of the natural numbers that reveal why (C) is true. And similarly, Bengson also appeals to this notion of a modal ability when he speaks of the intuition about the evil genius scenario as allowing one to "see" how 'thinking *guarantees* that one exists', and that the opposite would be 'impossible'. Thus, all such contemporary accounts define intuitions in terms of a mental state of grasp that is similar to that used to characterise the state of interrogative understanding.

In sum, both early and contemporary accounts of the nature of intuitions putatively suggest a link between intuitions and interrogative understanding. Again, the central motivation for this proposal is that these accounts appeal to a sense of 'grasping' or 'seeing' that is regarded as characteristic of the epistemic state of interrogative understanding. For instance, both Leibniz and Descartes make recourse to the idea that intuitions enable a grasp of their contents that is analogous to a grasp of material objects, which is similar to that which Hills and Grimm appeal to in order to describe the grasp involved in interrogative understanding. Furthermore, Bengson, Bonjour, and Chudnoff all appeal to the idea that intuitions bring into play a modal ability to 'see' why their contents are true, which is similar to the notion of 'grasp' that Grimm uses to characterise interrogative understanding.

For the remainder of this chapter, I take the view that veridical intuitions can enable interrogative understanding of their contents as a working hypothesis. In what follows, I seek to clarify this proposal and explore its implications for debates about the epistemology of intuitions. For this aim, I take on a two-step approach. I will first argue for a novel contrastive view of interrogative understanding—that is, the proposal that to understand why p is always to understand why p rather than some alternative q. I then build on this proposal to flesh out the idea that intuitions afford interrogative understanding of their contents.

5.2 Why Interrogative Understanding is Contrastive

5.2.1 Contrastivism

The set of proposals falling under the banner of 'contrastivism' have their roots in the idea first mooted by Dretske (1972) that certain statements naturally elicit a contrast between a set of relevant alternatives. An example proves helpful to begin elucidating this idea.[10] First, suppose you have three friends who have been variously misinformed about a recent transaction of yours. One of them asks you for an explanation as to why you gave your car to Peter. You then explain that

(A) I *sold* my car to Peter.

Your second friend then asks why you sold your car to Alice; you clarify that

(B) I sold my car to *Peter*.

[10] This example, as well as the one that follows, model the ones developed by Dretske (1972).

Finally, the third asks you why you sold your motorcycle to Peter. You then inform her that

(C) I sold my *car* to Peter.

Each of the statements (A)–(C) embodies a distinct kind of contrast. Statement (A) suggests a contrast between the alternatives of *giving* and *selling* the car to Peter. Statement (B) focuses on different alternatives for *who* bought your car (Alice or Peter). And in (C), the relevant contrast class concerns *what* was bought—a *car* or a *motorcycle*. Thus, superficially similar statements—such as (A)–(C)—can be said to plausibly elicit distinct *kinds* of contrast classes.

These humdrum observations gain notable theoretical weight once we recognise that differences in the contrasts embodied by superficially similar claims can be semantically significant—such that the particular set of alternatives that we focus on in interpreting a given statement can matter greatly for an evaluation of the truth-value of such statement. To clarify this proposal, consider how the statement

(D) Sue gave me the tickets by mistake.

can be either true or false, depending on which alternatives we focus on. For example, suppose that we are trying to identify what particular mistake Sue committed in this case. In this case, the semantic-value of this statement may depend on whether we read this as

(D') Sue gave *me* the tickets by mistake.

or

(D") Sue gave me *the tickets* by mistake.

Suppose that Susie was meant to give the tickets to Patrick but that she eventually hands them to me. In this case, (D') is true but (D") is false. The fact that the only difference between these two statements is the relevant set of alternatives we focus on (*who* the tickets were given to and *what* was given to me) shows that the particular contrast class is thus semantically significant.

The idea that statements can be contrast-sensitive in this way has often been used to motivate contrastive theories about a number of central philosophical concepts (for a collection Blaauw, 2012). Arguments for such contrastive views build up from the observation that claims employing those concepts naturally embody semantically significant contrasts between a particular set of alternatives. Perhaps the most well-accepted and prominent development of this approach is contrastivism about explanation.[11] On this view, whether some fact explains why *p* depends on which alternatives to *p* we focus on. For instance, suppose I cite the fact that Peter was hungy (E) as an explanation for why Peter ate the apple. Now, it is quite reasonable to think that it is true that 'E explains why Peter *ate* the apple (rather than just stay seated on the sofa)'; yet, it is intuitively false to say that 'E explains why Peter ate the *apple* (rather than the pear, or the pizza)', since these different options of food could might as easily have satiated Peter's hunger. In

[11] See van Fraassen (1980) for a detailed discussion of this proposal.

effect, this example provides reasons to think that whether some fact succeeds in explaining why Peter ate the apple depends on which particular class of contrasts we focus on—e.g., that between Peter's actions, or on his choices of food.

Other philosophical concepts that have been given similar contrastive treatments include (among others) causation, reasons, freedom, and knowledge.[12] Proponents of these accounts take on a similar strategy to that detailed above. First, they attempt to show how statements employing such concepts naturally suggest a contrast between a particular set of options. Then, they indicate that which specific contrast class we consider can influence whether we judge a statement using those concepts to be true or false. In addition, many contrastive accounts aim to show that contrastivism has significant theoretical advantages—helping to solve thorny philosophical puzzles.[13] For the remainder of this section, I argue that similar considerations motivate contrastivism about interrogative understanding.

5.2.2 Contrastivism about interrogative understanding

A first set of motivations for contrastivism about interrogative understanding are linguistic considerations. In particular, some cases our judgement of whether 'S understands why *p*' are sensitive to the particular alternatives to *p* we are considering. The following example is helpful to illustrate. Suppose that Anna has a good friend called Patrick, who she knows always has cake for dessert. Now, imagine that both are at a buffet and Anna sees Patrick choose chocolate cake. In this case:

(i) Anna understands why Patrick picked the chocolate cake rather than the crumble for dessert.

(ii) Anna understands why Patrick picked the chocolate cake rather than the lemon cake for dessert.

Many people take (i) to be true, yet they judge (ii) to be false. What is responsible for this shift in people's evaluations of Anna's epistemic states? One plausible explanation is that whether or not we take Anna to understand why Patrick chooses chocolate cake hinges on what other options of dessert we are considering. In particular, it depends on whether we contrast chocolate cake with other types of dessert (cake or crumble), or if we contrast this with different types of *cake* (lemon or chocolate). To bolster this explanation, here is another way of teasing out this shift in evaluations of Anna's epistemic states: I take it that readers will judge it to be true that

(ia) Anna understands why Patrick picked chocolate *cake* for dessert.

but that they will judge it to be false that

(iia) Anna understands why Patrick picked the *chocolate* cake for dessert.

Again, a plausible explanation for this shift in our judgements of Anna's epistemic states is that (ia) and (iia) elicit different dessert options from which Patrick could have chosen. Specifically, the emphasis on

[12] For contrastivism about causation and freedom, see Blaauw (2012); regarding reasons, see Snedegar (2017); and for contrastivism about knowledge, see Schaffer (2004).

[13] For instance, Schaffer (2007) has famously argued that contrastivism helps to diffuse the threat of scepticism, whilst maintaining closure.

cake in (ia) suggests a contrast between chocolate cake and other *kinds* of desserts. On the other hand, the emphasis in (iia) suggests that there were more choices of cake for Patrick to choose from. Since the only difference between (ia) and (iia) is the emphasis on different parts of the sentence, and that this plausibly causes a shift in the set of options they elicit, then plausibly this is what is responsible for people's shift in judgement between (ia) and (iia).

This pattern of variation in our evaluation of Anna's epistemic states are puzzling. After all, we typically interpret claims of the form 'S understands why *p*' as *outright* ascriptions of interrogative understanding. That is, we often take attributions of interrogative understanding to be invariant with respect to which alternatives to *p* we consider. Thus, if I claim to understand why Peter ate an apple (because he was hungry), presumably this is unaffected by whether or not I understand why he ate the banana or the pizza instead. But, the example of Anna detailed above provides a clear case where attributions of interrogative understanding are sensitive in just this way (such that our evaluations of whether or not a subject understands why *p* turns on which alternatives to *p* we entertain). In effect, this means that attributions of the form 'S understands why *p*' are not *outright* attributions of interrogative understanding, insofar as they can vary with respect to the particular alternative options to *p* we consider.

A second related line of support for contrastivism about interrogative understanding is that claims of the form 'S understands why *p*' are *question-relative*. That is to say that their truth-value changes with respect to the particular 'why *p*?' question they embed. To begin clarifying this proposal, recall that ascriptions of interrogative understanding are commonly taken to posit a relation between a subject and a correct answer to a question (sec. 5.1.1). Thus, the claim 'S understands why *p*' establishes an epistemic state of understanding between S and the answer to 'why *p*?'. Now, note that questions of the form 'why *p*?' naturally suggest a contrast between *p* and a set of alternatives insofar as they ask why *p* rather than some other occurrence obtains. Sometimes this contrast is made explicit: e.g., the question 'why did Peter eat the apple' can be meant as 'why did Peter eat the apple *rather than* eat the pear?', or alternatively as 'why did Peter eat the apple *rather than* just stay seated on the sofa?'. However, such contrasts may be implicit, and are only recovered by interpreting 'why *p*?' within a particular *context*—for instance, by attending to intonational stress, such as in 'why did Peter *eat* the apple?' (instead of just observing it) or 'why did Peter eat the *apple*?' (instead of the pear). Crucially, the correct answer to the relevant 'why *p*?' question hinges on which particular set of alternatives to *p* are under consideration. For instance, the appropriate answer to the question 'why did Peter eat the apple?' is not the same if we ask why he ate the apple instead of the pear (say, because he does not like pear) or if we ask why he ate that instead of remaining seated (e.g., because he was very hungry). Similarly, the answer for 'why did you *drive* to the store?' (because I was tired) can be very different from the answer to the question 'why did you walk to the *store*?' (because I needed to buy something).

Now, a natural consequence is that ascriptions of interrogative understanding should be expected to also depend on which particular 'why *p*' question we are considering. After all, attributions of interrogative understanding posit relations between subjects and answers to why-questions; and, if such answers vary depending on the particular why-question under consideration, then attributions of interrogative

understanding should be sensitive to these differences. Indeed, we can see this kind of question-relativity at work in the case of Anna (detailed above). As I have mentioned, most people find it quite natural to judge (i) to be true, yet judge (ii) as false. Another way to interpret this evaluation is that people judge Anna to understand why Patrick picks chocolate cake relative to the question 'why does Patrick pick chocolate cake rather than rhubarb crumble?', yet they deny that Anna understands Patrick's choices with respect to the question 'why does Patrick pick chocolate cake rather than lemon cake for dessert?'. Similarly, this question-relativity also explains the shift we find between evaluations of (ia) and (iia). These statements concern superficially similar why-questions (viz., 'why does Patrick pick chocolate cake for dessert?'). However, the different intonational stress in each of these statements suggest distinctive contrast classes (viz., different types of cake and different kinds of dessert), which yield distinct interpretations of this same question—and, plausibly, control our evaluation of each of these claims.

Another set of arguments in support of contrastivism about interrogative understanding is that this view has significant theoretical advantages. For instance, this view aptly explains the gradability of interrogative understanding ascriptions. This feature of our attributions of interrogative understanding is made evident in how a person might be said to understand why *p* to a *greater* or *lesser* degree than another. Thus, it is reasonable to expect that a 7-year-old child is likely understand to why the Earth orbits the Sun to a lesser degree than the common adult—who presumably is better acquainted with a greater number of relevant astronomical facts. More precisely, we can imagine that whilst the child may cite facts about the Sun's gravitational pull on Earth, and how this prevents the Earth from drifting off into space, this child might nevertheless fail to understand why this gravitational pull doesn't just cause the Earth to collide into the Sun. And insofar as the adult might understand why this is so—because she understands why the centripetal force from the Earth's rotation prevents it from colliding into the Sun—she can then reasonably be said to have fuller understanding of why the Earth orbits the Sun. Elaborating on this example in light of the suggestion that interrogative understanding is contrastive suggests the following description of the child's and the adult's respective epistemic states:

- The child understands why the Earth orbits the Sun rather than drift off into space.
- The adult understands why the Earth orbits the Sun rather than drift off into space, or collide into the Sun.

This way of fleshing out the epistemic states of the child and the adult provides a straightforward explanation for why they have differing degrees of understanding of why the Earth orbits the Sun. In more detail, the child can be said to understand why the Earth orbits the sun to a *lesser* degree because her epistemic state does not enable her to rule out that the Earth just collides into the Sun. And it is the adult's understanding of why this doesn't happen that explains why she can reasonably be attributed a greater degree of understanding of this topic. More generally, the above considerations suggest that the degree to which a person understands why *p* is a direct function of the number of alternatives to *p* she can be said to rule out. On this view, the greater the the set of contrasts of *p* that a subject S can eliminate, the more S can be said to understand why *p*. Relying on this account, we can also easily explain why an astrophysicist can be said to have a much greater degree of understanding of why the Earth orbits the Sun

when compared to the common adult. In particular, we can expect that the astrophysicist's expertise will enable her to eliminate a greater number of other alternatives, beyond those that the common adult is able to. That is, we can plausibly say that this astrophysicist understands why the Earth orbits the Sun rather than drift off into space, or collide into the Sun, or remain static, etc.

Another welcome theoretical upshot of contrastivism about interrogative understanding is that it helps to explain another notable pattern in our ascriptions of this epistemic state: viz., that our attributions of interrogative understanding vary across contexts. To illustrate, consider how we might reasonably claim that a 7-year-old child understands why the Earth orbits the Sun in a classroom setting. Suppose that this is because we assess that the child has indeed understood the influence of the gravitational pull from the Sun and can roughly explain how this determines the Earth's orbit around it. However, we might be expected to retract this attribution in the context of an astrophysics conference. Now, this doesn't mean that we thereby judge that the child doesn't understand *at all* why the Earth orbits the Sun. Rather, this retraction is due to our recognition that the minimal conditions to be attributed interrogative understanding of this fact in the latter context are much more demanding than those of the classroom setting.[14]

The contrastive account of interrogative understanding provides a straightforward explanation of this shift in our attributions. To clarify, recall that ascriptions of interrogative understanding posit an epistemic relation between a subject and a correct answer to a 'why p?' question. Now, as discussed earlier in this section, what determines whether a response to a 'why p?' question is appropriate will depend on how that question is interpreted in a particular context. Specifically, it will depend on which set of alternatives to p that are under consideration when that question is posed. For instance, asking 'why does the Earth orbit the Sun?' in a classroom of 7-year-olds after teaching them about the gravitational pull of the Sun on the Earth makes salient only a restricted number of relevant contrasts to this fact. In particular, it underscores the contrast between the Earth orbiting the Sun rather than drifting off into space. Thus, the correct answer to this question in the classroom setting will be that which rules out this particular contrast to the relevant fact. And to the extent that the child can be said to stand in the appropriate epistemic relation to *this* answer, she can reasonably be said to understand why the Earth orbits the Sun. However, note that asking this superficially same question within the context of an astrophysics conference plausibly brings to light a much greater number of relevant alternatives to the fact that the Earth orbits the Sun. For instance, asking this in a panel of specialists on the gravitational pull of the Sun makes salient the contrast between the Earth orbiting the Sun rather than, say, drifting off into space, colliding into the Sun, having a star-shaped orbit, remaining static, etc. In this sense, we can assume that the correct answer to this question when posed in the context of a astrophysics conference will be one that rules out these many more alternatives. And, given that the child does not possess an answer that rules out these many alternatives, then she can reasonably be said to not understand why the Earth orbits the Sun. Thus, our shift in attributions of interrogative understanding is naturally explained as a function of what is taken to constitute the correct answer to the relevant "why p?" question in each particular context—which itself depends on the alternatives to p that are made salient when that question is posed.

[14] See, e.g., Hills (2015) and Sliwa (2017).

5.2.3 Taking Stock

The observations in this section offer good reasons in support of contrastivism about interrogative understanding—i.e., the proposal that to understand why *p* is to understand why *p* relative to some particular set of alternatives. In more detail, I first motivated the idea by way of linguistic considerations. First, that our evaluation of the claim 'S understands why *p*' may vary depending on which alternatives to *p* we consider; and second that attributions of interrogative understanding are *question-relative*. I then argued that this view has two significant theoretical advantages. First, it naturally captures the widely-shared intuition that interrogative understanding is gradable. Second, this proposal provides resources with which to explain the mechanisms governing a notable shift in our attributions of interrogative understanding across different contexts.

Now, despite providing a moderate degree of support for contrastivism about interrogative understanding, the above considerations do not entail this view. Indeed, it is still very much an open possibility to accept all the data motivating contrastivism about interrogative understanding, yet deny that contrastivism provides the best explanation of this data. One of the most promising ways of fleshing out this proposal is to give an adequate semantic account of this contrast-sensitivity that does not make recourse to the idea that such statements are actually contrastive. Alternatively, one could attempt to offer a pragmatic explanation that accounts for the contrast-sensitivity of interrogative understanding attributions as a mere feature of conversational propriety that does not provide any fundamental insights into nature of this epistemic state. In the next section, I develop and reject the most promising ways to flesh out these alternative approaches to explaining the contrast-sensitivity of interrogative understanding ascriptions.

5.3 Non-Contrastive Alternatives

5.3.1 The Conjunction Strategy

The Conjunction Strategy (CS) is perhaps the most natural way to push back against contrastivism about interrogative understanding.[15] According to this view, occurrences of "rather than" clauses in ascriptions of interrogative understanding should be read as "and not". Thus, the (CS) amounts to the claim that the object of understanding is a *negated conjunct* of the relevant contrasts: to understand why *p* rather than *q* just is to understand why *p* **and not**-q. More schematically:

(CS) S understands why A rather than B iff S understands why A **and not**-B.

(CS) neatly captures the kinds of examples that appear to motivate contrastivism about interrogative understanding. To see how, consider again the case of Anna, who knows that Patrick always has cake for dessert. In this scenario, the following ascriptions of epistemic states seem apt:

(i) Anna understands why Patrick picked the chocolate *cake* rather than rhubarb *crumble* for dessert.

[15] Both Schaffer (2008) and Snedegar (2017) discuss the conjunction strategy as plausible (yet ultimately untenable) alternatives to contrastivism.

(ii) Anna does not understand why Patrick picked the *chocolate* cake rather than *lemon* cake for dessert.

According to the (CS), (i) boils down to the claim that Anna understands why Patrick chooses chocolate cake **and not** the crumble. This reading is plausible: given that Anna knows that (E)—i.e., that Patrick always to chooses cake—it is rather easy to see why she understands why Patrick would choose chocolate cake and not crumble. Furthermore, (CS) also nicely explains why people tend to judge that Anna *doesn't* understand why Patrick would choose the chocolate cake rather than the lemon cake: her knowledge that (E) doesn't seem sufficient to explain why Patrick would choose the chocolate cake **and not** the lemon cake. So, to a first approximation, it seems as if (CS) can account for the kinds of cases motivating contrastivism about interrogative understanding without recourse to the idea that they are contrastive. However, this strategy has some significant shortcomings. To explain, let us focus on a different example.

Suppose Barbara and Quinn are friends, and that (E') Quinn is trying to lose weight. Now imagine Barbara knows that (E'), and that she is wondering about Quinn's new habits. The following two ascriptions of epistemic states seem correct:

(iii) Barbara understands why Quinn has a low-fat yoghurt rather than a chocolate brownie after his meal.

(iv) Barbara understands why Quinn has tea rather than a low-fat yoghurt after his meal.

According to (CS), the correct reading of these ascriptions would be

(iiia) Barbara understands why Quinn has a low-fat yoghurt and not a chocolate brownie after his meal.

(iva) Barbara understands why Quinn has tea and not a low-fat yoghurt after his meal.

The problem for proponents of (CS) is that for this view to provide a genuine non-contrastive account of the contrast-sensitivity of statements, it must be committed to the idea that to understand why *p* rather than *q just is* to understand why *p and* to understand why **not**-q. That is, (CS) has the following corollary:

(CS') S understands why A rather than B iff S understands why A and understands why not-B.

Some might seek to deny that (CS) entails (CS'). In effect, this response denies that interrogative understanding distributes over the relevant contrasts. On this alternative reading, the (CS) would amount to the claim that the object of understanding is the negated conjunct 'p-and-not-q'. That is, to understand *p* rather than *q* just is to understand p-and-not-q. On this view the (CS) entails that subjects don't understand why *p simpliciter*; rather, subjects only understand why *p* relative to some set of contrasts. But, note that this reading of (CS) is in effect a contrastive view of interrogative understanding: it essentially implements another way of incorporating the relevant contrast in the object of understanding. As such, I will reject this alternative reading of (CS).

(CS') raises a challenge for proponents of the (CS), insofar as it commits them to saying that

(iiib) Barbara understands why Quinn has a low-fat yoghurt after his meal.

Furthermore, according to (CS'), (iva) entails that

(ivb) Barbara understands why Quinn does not have a low-fat yoghurt after his meal.

The challenge for a proponent of (CS) is that this account seems insufficient to explain the shift between (iiib) and (ivb). Without recourse to the idea that these statements embody a contrast class, it is difficult to see how Barbara can both understand why Quinn has and does not have a low-fat yoghurt As such, the (CS) lacks the resources to explain the contrast-sensitivity we find in ascriptions of interrogative understanding.

5.3.2 Degrees of Understanding Strategy

Another natural way to push back against contrastivism about interrogative understanding is to read the "rather than" clauses to mean "more than". On this view, the contrast brought out by these clauses amounts to a straightforward comparison in how much a subject understands *each* of the relevant contrasts. In other words, to say that S understands why A rather than B *just is* to say that S understands why A *more than* she understands why B—I term this the Degrees of Understanding Strategy (DUS). More schematically:

(DUS) S understands why A rather than B iff S understands A *more than* why B.

(DUS) builds on the natural suggestion that interrogative understanding comes in degrees. However, it is noteworthy that if the (DUS) is to get off the ground, it will require a particular elaboration of the idea that interrogative understanding comes in degrees. Specifically, the (DUS) relies on the idea that even in cases where a subject S does not understand why *p*, they nonetheless do *understand* why *p* to a nil degree. Given this reading, the (DUS) can then be said to account for the type of case that seems to motivate contrastivism about interrogative understanding. For instance, recall that in the case of Anna we considered the following attributions to be plausible:

(i) Anna understands why Patrick picked the chocolate *cake* rather than rhubarb *crumble*.

(ii) Anna does not understand why Patrick picked the *chocolate* cake rather than *lemon* cake.

In line with the (DUS), we can read (i) to say that Anna understands why Patrick picks chocolate cake *more than* she understands why Patrick picks rhubarb crumble. This seems plausible: since Anna does not understand the latter *at all*, then she clearly understands Patrick's choice of chocolate cake to a greater degree. Furthermore, on the (DUS), (ii) can be read as saying that Anna does not understand why Patrick picks chocolate cake *more than* she understands why he picks lemon cake. Again, this seems right: if Anna has *no* understanding of why Patrick would choose chocolate cake instead of lemon cake, then it would seem right to say that she fails to understand both of these things.

The (DUS) can also account for the kind of examples that were shown to raise trouble for the (CS). For instance, recall the case of Barbara and Quinn. Given that Barbara knows that (E')—i.e., that Quinn is trying to lose weight—then the following epistemic states seem right:

(iii) Barbara understands why Quinn has a low-fat yoghurt rather than a chocolate brownie after his meals.

(iv) Barbara understands why Quinn has tea rather than a low-fat yoghurt after his meals.

According to (DUS), these statements could be read as:

(iiic) Barbara understands why Quinn has a low-fat yoghurt after his meals more than why he has a chocolate brownie after his meals.

(ivc) Barbara understands why Quinn has tea after his meals, more than why he has a low-fat yoghurt after his meals.

What is crucial to note here is that the proponent of the (DUS) is not committed to the kind of tricky entailments that cause trouble for the proponent of (CS). Specifically, she need not ascribe interrogative understanding of the negation of one of the contrasts, and is not thereby committed to saying that Barbara both understands why Quinn would choose and would **not** choose a low-fat yoghurt. Instead, the (DUS) interprets (iii) and (iv) purely in terms of the degree to which Barbara understands each of the relevant contrasts. As such, we get the result that Barbara understands why Quinn has a low-fat yoghurt after his meals *more than* why he has chocolate brownie after his meals, yet *less than* why he has tea.

Despite this moderate degree of support for the (DUS), this view has a notable shortcoming. Specifically, this proposal runs into trouble because it establishes a *strict* relation between the contrast-sensitivity we find in ascriptions of interrogative understanding and differences in *degrees* of understanding. Now, a natural consequence of this proposal is then that the negation of 'S understands why p rather than q' either implies that S has a *same degree* of interrogative understanding of both p and q, or that S actually understands why q more than she understands why p. To see why this strict relation between contrast-sensitivity and degrees of understanding causes trouble for the DUS, consider the following example.

Anna and Patrick are making predictions about the performance of the soy industry this year. Suppose, for the sake of argument, that both of them have a same degree of understanding of why this sector of the economy will *remain stable*. But, let's imagine that they have differing pieces of information leading to this conclusion:

E1 Patrick learns that the government didn't provide *additional* subsidies to the soy industry this year, and so this sector of the economy will not grow as previously expected.

E2 Anna learns that despite the sharp decrease in China's demand for soy, Korea's demand for the grain has soared, and so this sector of the economy will not slump as previously expected.

Given this scenario, the following then seems right:

(v) Patrick understands why the soy industry will remain stable rather than grow.

(vi) Anna understands why the soy industry will remain stable rather than slump.

Now, according to the (DUS), we should interpret (v) and (vi) to read:

(va) Patrick understands (to degree d1) why the soy industry will remain stable more than (to degree d2) why it will grow. (d1>d2)

(via) Anna understands (to degree d1) why the soy industry will remain stable more than (to degree d3) why it will slump. (d1>d3)

So far so good. But, now consider the following attribution of epistemic states:

(vii) Patrick doesn't understand why the soy industry will remain stable rather than slump.

(viii) Anna doesn't understand why the soy industry will remain stable rather than grow.

I contend that both (vii) and (viii) are true. After all, Patrick only has enough information to provide him with an understanding of the soy industry's stability when compared to a possible scenario of this sector's growth—but, not with respect to the situation where it would slump. Similarly, Anna understands why the soy industry will remain stable with respect to the possibility of it slumping, but not of its potential growth. The challenge for a proponent of the (DUS) is that on this view (vii) and (viii) are to be read either as:

(viia) Patrick understands (to degree d1) why the soy industry will remain stable to a *same degree* (to degree d1) as to why it will slump.

(viiia) Anna understands (to degree d1) why the soy industry will remain stable to a *same degree* (to degree d1) as to why it will grow.

or

(viib) Patrick understands (to degree d4) why the soy industry will slump *more than* he understands (to degree d1) why it will remain stable (d4>d1).

(viiib) Anna understands why the soy industry will grow (to degree d5) *more than* she understands (to degree d1) why it will remain stable (d5>d1).

These readings sound implausible; the correct explanation for (vii) and (viii) seems to be instead that Patrick and Anna have a restricted amount of information about the soy industry's performance, and as such they are not in a sufficiently good epistemic position to understand its stability with respect to every potential scenario (growth/slump). The example above thus underscores the inability of the (DUS) to capture this subtlety: in particular, it shows how there are distinct *ways* of understanding the performance of a given sector of the economy and that these make salient incompatible sets of contrasts (stable rather than grow, or stable rather than slump). Thus, ascriptions of interrogative understanding are sensitive also to the *quality* of the information, and to the *kind* of interrogative understanding that this enables. By forcing a reading of "rather than" clauses purely in terms of degrees of understanding, the (DUS) fails to capture this pattern in our ascriptions of interrogative understanding. In light of this difficulty, I propose that the (DUS) fails.

5.3.3 Pragmatic Explanation

We have now considered and rejected two (non-contrastive) semantic accounts of the contrast-sensitivity we identified in claims the form 'S understands why p'. These arguments provide grist for the mill of contrastivism about interrogative understanding by suggesting that it provides the best semantic account of this contrast-sensitivity. However, these patent shortcomings of the (DUS) and the (CS) do not immediately *entail* that contrastivism holds. For instance, I have not yet considered the option that there could be a *pragmatic explanation* for these patterns of contrast-sensitivity—such that they reflect what is conversationally appropriate to say, but not what is necessarily *true*. In this section, I develop and reject one such account that closely models a prominent pragmatic explanation for the contrast-sensitivity of knowledge ascriptions. I first describe motivations for contrastivism about knowledge and detail the pragmatic explanation that has been raised against them. I then proceed to show how similar considerations do not naturally carry over to contrastivism about interrogative understanding.

Contrastive theories about knowledge are usually motivated by examples of the following type:

> **Mary the Detective** Last night, Peter robbed the jewelry store. He smashed the window, forced open the locked safe, and stole the rubies inside. But Peter forgot to wear gloves. He also forgot about the security camera.
>
> Today, Mary the detective has been called to the scene to investigate. So far she has the following evidence. She has been told that there was a theft, she has found and identified Peter's fingerprints on the safe, and she has seen and recognized Peter on the security video, filmed in the act of forcing open the safe. She has no further information.
>
> (Schaffer and Knobe, 2012, p. 689)

Most people who consider the case above of Mary the detective judge it to be true that

(ix) Mary now knows that Peter rather than anyone else stole the rubies.

yet, many people who read the above case judge it to be false that

(x) Mary now knows that Peter stole the rubies rather than anything else.

This pattern of evaluations raises a question as to why people judge these two statements differently. Contrastivists claim that people's intuitions about (ix) and (x) pick up on the intrinsic contrast-sensitivity of the relevant knowledge attributions—viz., between the contrast in who stole the rubies (Peter or someone else in (ix)) and in what was stolen (the rubies or something else in (x)). On the other hand, opponents of contrastivism about knowledge contend that—although widely shared—these intuitions are actually *false positives*. Gerken (2017) develops this objection in great detail. In this sense, he offers the following further reflection on the case of Mary the detective:

M1 In every case, Mary knows that Peter rather than anyone else stole the rubies only if Mary knows that Peter stole the rubies.

M2 In every case, Mary knows that Peter stole the rubies only if Mary is in a position to know that the rubies were stolen.

M3 In the present case, Mary is not in a position to know that the rubies were stolen.

M4 In the present case, Mary does not know that Peter stole the rubies.

M5 In the present case, Mary does not know that Peter rather than anyone else stole the rubies.

Gerken suggests that the above line of reasoning reveals the evidential demands for knowing that Peter stole the rubies to be higher than those for knowing that Peter stole *something*. Given that Mary does not have sufficient evidence to determine *what* was stolen—indeed, she even lacks relevant beliefs about the object of the theft (viz., rubies)—then, contrary to people's intuitions about (ix), it appears that Mary *does not know* that Peter stole the rubies after all (for similar views: DeRose, 2011; Kelp, 2011). This then raises the question as to why people are inclined to evaluate (ix) to be true, given that M1–M5 provides very clear reasons to deny this evaluation. One pragmatic explanation is that people mistakenly evaluate (ix) as true because this is conversationally appropriate, albeit false.

One promising approach to fleshing out this suggestion builds on the well-developed body of work on presupposition accommodation. The latter refers to a process whereby speaker/hearers update the common-ground in order to adjust to the presuppositions of the many moves within communication. For instance, most people who hear that 'John regrets that Sally didn't call him' will accept the presupposition that Sally didn't call John. And similarly, most will take it from the utterance 'It was John who stole the bike' that at least some bike was stolen. Building on this familiar set of observations, Steglich-Peterson (2015) proposes a pragmatic interpretation of the data regarding people's evaluations of (ix) and (x). Specifically, he argues that (ix) triggers the *false* presupposition that Mary knows *what* was stolen, which speaker/hearers then promptly grant and include as part of the common-ground. In this sense, the judgement of (ix), according to which Mary knows that the rubies were stolen is simply a feature of *conversational propriety*—in line with the natural process of presupposition accommodation—but, which is ultimately false.

The cogency of this kind of pragmatic explanation for contrastivism about knowledge raises the question of whether a similar line of reasoning can challenge contrastivism about interrogative understanding. That is, it is quite natural to wonder whether contrastivism about interrogative understanding may also be motivated by intuitions which, although conversationally appropriate, are nonetheless ultimately false. As an attempt at examining whether this position is tenable, consider the following case:

> **Hector from HR** One of Hector's duties as director of HR is overseeing the annual process of reshuffling, whereby some employees are either reassigned or fired. Hector's responsibility in this position is to gather records of all those employees who will be part of the reshuffling; but, he does not take part in the process that determines who gets fired or reassigned (which happens behind closed doors).

> Hector knows that the only employee selected for this year's reshuffling is Susan. He later learns that Susan was fired.

After considering this case, I expect that the reader will judge it to be true that

(xi) Hector understands why *Susan rather than any other employee* was fired.

but, that the following will seem false:

(xii) Hector understands why Susan was *fired rather than reassigned*.

This pattern of evaluations bears important parallels with the judgements elicited by the case of Mary the detective. Recall that in response to the latter case, most people are inclined to judge that Mary knows that *Peter rather than anyone else* stole rubies, but deny that Mary knows that Peter *stole the ruby rather than something else*. Similarly, I expect that most readers will judge that Hector understands why *Susan rather than anyone else* was fired, yet I expect that most readers will also judge that Hector does not understand why Susan was *fired rather than reassigned*. Critically, this raises the suspicion that, as has been argued for the case of Mary the Detective, perhaps the intuitions regarding (xi) and (xii) are also false positives. In particular, it suggests a line of reasoning analogous to M1–M5 can be drawn for the case of Hector from HR:

H1 In every case, Hector understands why Susan rather than anyone else was fired only if Hector understands why Susan was fired.
H2 In every case, Hector understands why Susan was fired only if Hector is in a position to understand why a particular employee was fired.
H3 In the present case, Hector is not in a position to understand why any particular employee was fired.
H4 In the present case, Hector does not understand why Susan was fired.
H5 In the present case, Hector does not understand why Susan rather than any other employee was fired.

If we assume that this line of reasoning is accurate, then this raises the question as to why we are inclined to judge that (xi) is true despite its explicit conflict with H5. Here we can offer a pragmatic explanation that, like that advanced for contrastivism about knowledge, also builds on the notion of presupposition accommodation. Thus, it can be suggested that when people read 'Hector understands why Susan rather than any one else was fired', they accommodate the false presupposition that Hector is in a position to understand why a particular employee was fired. This then explains why judging (xi) to be true is conversationally appropriate, despite being ultimately false.

Although seemingly compelling, I propose that there are good reasons to resist this account. Specifically, I contend that H1–H5 fails insofar as H3 begs the question against contrastivism about interrogative understanding. To clarify, it is helpful to note how H3 is itself contrast-sensitive. For instance, the following reading of H3 is true:

H3' In the present case, Hector is not in a position to understand why any particular employee was *fired rather than reassigned.*

After all, it is quite clear that Hector is not in a position to understand why employees get fired rather than reassigned: after all, he is not part of the decision-process that discusses this, and therefore he is not privy of the reasons for which employees either get fired or reassigned. But note that the following reading of H3 is false:

H3" In the present case, Hector is not in a position to understand why any particular employee was *fired rather than remain in their current position*.

The explanation for this is that Hector is indeed in a position to understand why employees get fired rather than remain in their current positions: viz., he knows which employees are up for reshuffling, and thereby which ones are liable to be fired, rather than remain in their current positions.

The central upshot of these considerations is that this pragmatic explanation for contrastivism about interrogative understanding is unmotivated. Recall that this theory seeks to explain how our intuitions in response to the case of Hector from HR are conversationally appropriate, yet ultimately false. In more detail, this theory suggests that our intuition that (xi) is true is due to a natural process of presupposition accommodation—by which we wrongly accept the false presupposition that Hector is in a position to understand why any particular employee gets fired. I contend that this presupposition is not false because constructions of the form 'S is in a position to understand why p' are contrast-sensitive. As pointed out above, we can see this contrast-sensitivity at work in the different readings of H3. On this note, I have suggested that H1–H5 is false. Ultimately, this undermines what I consider to be the most promising and powerful pragmatic explanation for contrastivism about interrogative understanding.

5.3.4 Summing up

I have been arguing against alternative (non-contrastive) accounts of the data motivating contrastivism about interrogative understanding. First, I showed that the Conjunction Strategy (CS) lacks the resources to explain the kind of contrast-sensitivity we find in certain attributions of interrogative understanding. Second, I argued that the Degrees of Understanding Strategy (DUS) gives the wrong result in a set of very tricky cases of interrogative understanding-ascriptions. And lastly, I detailed why the most promising pragmatic explanation of this data is ultimately unmotivated. Given that these alternative views face significant shortcomings in accounting for the contrast-sensitivity of attributions of interrogative understanding, I then propose that contrastivism is to be preferred. Thus, in line with arguments for contrastivism about other concepts in the philosophical literature,[16] I suggest that interrogative understanding is contrastive.

[16] See Blaauw (2012).

5.4 Intuitions and Contrastivism about Interrogative Understanding

In the outset of this chapter, I discussed the view that intuitions can enable interrogative understanding of their contents. In this section, I tease out the implications of contrastivism about interrogative understanding for this proposal. First, I explain how focus on the contrastive nature of interrogative understanding helps to better characterise the epistemic state intuitions afford with respect to their contents. Second, I detail how this proves useful to precisify the warrant intuitions are poised to provide to philosophical positions and claims.

To develop the arguments in this section, I will focus on a set of intuitions about the so-called Trolley problem—i.e., a set of hypothetical scenarios where people are asked if it would be morally permissible to alter the course of a run-away trolley to save many people at the cost of sacrificing the life of one person (Thomson, 1985). To begin, consider the following well-known variation of this scenario:

Lever: A runaway trolley is headed toward five innocent people who are on the track and who will be killed unless something is done. Patrick may pull a lever, which will redirect the trolley onto a second track, saving the five people. However, on this second track is an innocent bystander (Susy), who will be killed if the trolley is turned onto this track.[17]

It is well-documented that many people who read **Lever** have an intuition that it would be morally permissible for Patrick (or anyone in his position) to pull the lever in this case—thus sacrificing one person to save the lives of five others. Moreover, a great number of philosophers take this intuition to be somehow representative of moral truth.[18] Thus, they contend that it would indeed be morally permissible for one to pull the lever in this case. In effect, this amounts to the suggestion that this intuition about **Lever** is veridical. For current purposes, I will also assume that this is true. Furthermore, in line with the idea that veridical intuitions enable interrogative understanding of their contents, I will assume that intuitions about **Lever** provide understanding of *why* it is morally permissible to pull the lever and sacrifice one person to save the life of five others. Now, in previous sections I defended the idea that interrogative understanding is contrastive. In what follows, I argue that focus on the contrastive nature of interrogative understanding proves instructive to better characterise the nature of the epistemic state that this intuition about **Lever** affords with respect to its content.

First, recall that the contrastive view I defended in previous sections amounts to the claim that to understand why p is to understand why p rather than some alternative, q. This proposal builds on the widespread idea that to understand why p is to stand in an epistemic relation to a correct answer to the question 'why p?'. But, as previously pointed out, what counts as the correct answer to 'why p?' will vary depending on how we interpret this question. For instance, whether an answer to 'why did Peter eat the apple?' is correct will depend on whether we are asking about Peter's choice of food (why he ate the *apple* rather than, say,

[17] This formulation of the Trolley problem is adapted from Liao et al., (2012).
[18] For a review: Kamm, (2015)

the pizza), or his actions (why did he *eat* rather than, say, hand it back to the customer). For instance, it would not suffice to explain Peter's odd action of taking the apple from the customer's hand and eating it by answering that he prefers apple to pizza. In this sense, I proposed that these considerations show how specifying the epistemic state ascribed by 'S understands why p' will require defining what is the particular reading of the relevant 'why p?' question being considered.

Now, let's return to the intuition that in **Lever** one would be morally permitted to pull the lever. Given this intuition is veridical and thus enables interrogative understanding of its contents, then it establishes an epistemic relation between a subject and a correct answer to the question 'why is it morally permissible to pull the lever to sacrifice one person and save the lives of five others?'. But, as just pointed out, what counts as a correct answer to this question will hinge on the particular reading of this question under consideration. For current purposes, it will be useful to clarify what is the particular reading of this question at stake in intuitions about **Lever**. I contend that in having this intuition, one will stand in an epistemic relation to the answer to 'why is it morally *permissible*, rather than *impermissible* to pull the lever to sacrifice one person and save the lives of five others?'. Otherwise put, the reading of the relevant question at issue here is that which contrasts the moral permissibility of the particular act of pulling the lever to sacrifice one person and save the lives of five others. This clarification concerning the intuition about **Lever** proves instructive in at least two respects. I detail each in turn below.

First, narrowing down on an answer to this particular reading of the relevant why-question thus helps to define what epistemic state is involved in having this intuition. For, note that this specification makes it clear that an intuition about **Lever** puts one in a privileged epistemic position with respect to two types of consideration: first, the claim that it is morally permissible to pull the lever in this case; and second, to a set of considerations that rule out *some* alternatives to this claim. Which alternatives? Those that suggest it would be impermissible to pull the lever. As such, the epistemic state involved in this intuition puts one in a privileged epistemic position to a set of considerations that provide reasons for why it is morally permissible to pull the lever.

A second way in which the above considerations prove instructive is that they help to clarify what kind of epistemic state is *not* involved in having an intuition about **Lever**. This is an important claim; but, to elucidate it, we must make a small detour. In particular, it will be useful to consider another well-known formulation of the Trolley problem:

Push: A runaway trolley is headed toward five innocent people who are on the track and who will be killed unless something is done. Patrick can push an innocent bystander (Susy) in front of the trolley. The runaway trolley would be stopped by hitting (and killing) Susy, but would thereby save the five people on the track.[19]

Most people who consider the scenario described in **Push** report having the intuition that it would be morally *impermissible* for Patrick to push the innocent bystander to sacrifice one person and save the lives

[19] This formulation of the Trolley problem is also adapted from Liao et al., (2012).

of five others. Now, some philosophers regard this intuition to be veridical. That is, they regard it to be a moral fact that it is impermissible to push the innocent bystander in this case. Now, we seem to have a conflict: how can we square the proposal that intuitions about **Push** are veridical whilst maintaining that intuitions about **Lever** are also veridical? After all, they seem to provide epistemic support to distinct claims: on one, that it is morally permissible to sacrifice one person to save five others; on the other, that it is impermissible to do so. I propose that focusing on the contrastive nature of interrogative understanding proves instructive here.

As I suggested above, an intuition about **Lever** will enable understanding of its contents. In more detail, I proposed that this intuition establishes an epistemic relation between a subject and the correct answer to 'why is it morally permissible rather than *impermissible* to pull the lever to sacrifice one person and save five others?'. As such, this intuition provides understanding of why it is permissible rather than impermissible to pull the lever to bring about this result. However, I contend that the intuition about **Lever** does *not* establish an epistemic relation between a subject and the answer to 'why is it morally permissible to pull the lever rather than *push an innocent bystander* to sacrifice one person and save five others?'. As such, an intuition about **Lever** will *not* put one in a privileged epistemic position with respect to those reasons that explain why it is morally permissible to pull the lever rather than push the innocent bystander onto the tracks.

These considerations help to precisify the kind of epistemic state one is in after having the intuition that it would be morally permissible to pull the lever. In particular, it would be a situation in which one is in a privileged epistemic standing with respect to considerations that rule out *some* (but not all) alternatives to the claim 'it is morally permissible to pull the lever to sacrifice one person and save five others'. In effect, these considerations show how the epistemic state afforded by the intuition to **Lever** is not in conflict with that afforded by the intuition to **Push**. The apparent conflict here would be analogous to that of understanding why Peter ate the apple rather than the pear (because he loves apples) but not understanding why Peter ate the apple rather than return it to the customer. In this case, one can explain why Peter ate the apple to some extent; but, one cannot explain it fully. Similarly, in the case where a subject S who has an intuition about **Lever**, she can explain why it is morally permissible to pull the lever to some extent; but, she cannot explain it fully. The epistemic state S stands with respect to this claim is then one wherein she can rule out *some* (but not every) alternative to this claim. Specifically, she cannot rule out the relevant alternatives regarding the particular action that leads to sacrificing one person to save the lives of five others.

A last noteworthy point is that these considerations have important implications for questions and debates about the methodology of philosophy. In particular, they prove useful to pinpoint *which* philosophical positions and claims an intuition provides warrant for. To clarify, let us focus on the above discussion about the epistemic state involved in having an intuition about **Lever**. As proposed above, this intuition establishes an epistemic relation to the answer of 'why is it morally permissible rather than impermissible to pull the lever and save the lives of five other people?' And, critically, I have suggested that this puts one in a privileged epistemic position with respect to the claim that it is morally permissible to do so, and

to considerations that rule out *some* alternatives to this claim. Which alternatives? Those that rule out why it would be morally impermissible to carry out this action. In this sense, I suggest that an intuition about **Lever** warrants considerations that rule out *those* alternatives. That is, they provide epistemic support to philosophical positions and claims that articulate the reasons for why it is morally permissible to *pull the lever* to sacrifice the life of one person and save five others. However, in line with the considerations about, this intuition does not provide warrant for considerations regarding the moral permissibility of *other* actions to bring about this same result. And, as such, an intuition about **Lever** does not thereby warrant philosophical positions and claims that articulate reasons for why it would be morally permissible to push the innocent bystander instead. These considerations thus suggest that by focusing on the contrastive nature of interrogative understanding, we can gain explanatory purchase on how intuitions provide warrant for philosophical positions and claims. That is because contrastivism allows for a better characterisation of the epistemic state an intuition affords with respect to its contents. On this proposal, a veridical intuition that p establishes an epistemic relation between a subject and p, as well as considerations that rule out *some* (but not all) alternatives to p. In this sense, that intuition will provide epistemic warrant for *those* considerations.

In sum, I propose that contrastivism about interrogative understanding helps to flesh out the claim that an intuition that p enables understanding of why p. First, I explained why this allows us to precisify the kind of epistemic state involved in having a particular intuition. On this proposal, an intuition that p affords a privileged epistemic standing with respect to both p, as well as considerations that rule out *some* (but not necessarily all) alternatives to p. Second, I argued that this elaboration provides insight into the warrant intuitions are poised to provide to philosophical positions and claims.

5.5 Conclusion

We have been exploring the idea that veridical intuitions enable interrogative understanding of their contents. To develop this proposal, I first argued for the novel thesis that interrogative understanding is contrastive. I first motivated this proposal by showing how statements of the form 'S understands why p' are contrast-sensitive. I then considered and rejected non-contrastive (semantic and pragmatic) accounts of this data. Given that these alternative views face significant shortcomings in accounting for such patterns of contrast-sensitivity, I proposed that contrastivism is to be preferred. Lastly, I argued that focus on the contrastive nature of interrogative understanding shows that a veridical intuition that p affords one a privileged epistemic standing with respect to p, and to considerations that rule out *some* alternatives to p. I then argued that this proves instructive to precisify what positions and claims an intuition can provide epistemic warrant for. These arguments thus underscore how inquiry into the nature of interrogative understanding proves useful to advance debates about the epistemology of intuitions, and about the epistemic role they play in philosophical inquiry.

Conclusion

We have been examining whether appeals to intuition provide warrant for philosophical positions and claims. To address this question, we have surveyed and assessed the main positions in recent debates surrounding the experimentalist challenge. Against Defenders, I argued that their attempts to undermine the experimentalist challenge fail. And against Critics, I proposed that empirical findings do not warrant wholesale scepticism about the use of intuitions in philosophy.

In chapter 1, I argued against a reply to the experimentalist challenge that builds on the thesis that philosophers do not rely on intuitions as evidence. According to this line of reply, if intuitions do not play an evidential role in philosophy, then empirical findings about intuition are irrelevant to assess the methodology of philosophy. Against this view, I argued that even if philosophers do not rely on intuitions as evidence, this leaves it open that intuitions play some other epistemically significant role in philosophy. I elaborated on this proposal to explain why, *pace* some Defenders, empirical findings about intuitions are relevant for assessment of the methodology of philosophy.

In chapters 2 and 3, I examined another prominent line of reply to the experimentalist challenge. On this proposal, empirical studies about intuition need to correct for glaring methodological flaws before they can be put to use in methodological debates. I argued that attempts to establish this proposal via an endorsement of a Phenomenalist conception of intuitions fail by appealing to findings on mental states with a feeling of rightness. I then showed how findings on the workings of these mental states contribute to improving traditional methods of philosophy. In the subsequent chapter, I examined an alternative formulation of this line of reply, according to which the relevant empirical studies rely on a problematic inference: namely, that the deleterious effects on non-philosophers' intuitions are representative of problems in the intuitions of professional philosophers. I argued that in at least some cases, the intuitions of philosophers are shielded from the biases afflicting the intuitions of non-philosophers. But, I explained that more empirical work needs to be done to establish what those cases are.

In chapter 4, I argued against Critics that nascent international studies testing for cross-cultural variation in intuitions would not lend weight to a radical version of the experimentalist challenge. To develop this argument, I outlined the epistemological considerations driving this proposal and teased apart two formulations of the relevant challenge. I then explained why both these formulations rely on controversial assumptions. In this sense, I proposed that even if an intuition is found to vary cross-culturally, this does not mean that it is thereby unsuitable for use in philosophical inquiry.

In chapter 5, I explored and developed the proposal that intuitions provide interrogative understanding of their contents. My approach was two-fold. First, I argued for a novel contrastive view of interrogative understanding—i.e., the view that to understand why p is to understand why p rather than some alternative, q. Then, I argued that focusing on the contrastive nature of interrogative understanding is useful to explain why intuitions are poised provide warrant for philosophical positions and claims.

Taken together, these arguments show that a moderate version of the experimentalist challenge enjoys considerable empirical and theoretical support. On this proposal, intuitions can provide warrant for philosophical positions and claims; but, philosophers should make recourse to experimental resources to determine when to trust their intuitions, and when to refrain from relying on them in inquiry. The arguments in this thesis thus demonstrate that experimental and traditional methods of philosophy can go hand-in-hand.

Bibliography

Ackerman, R and H Zalmanov (2012). "The persistence of the fluency-confidence association in problem solving". In: *Psychonomic Bulletin & Review* 19.6, pp. 1189–1192.

Ackerman, R. and V. A. Thompson (2017). "Meta-Reasoning: Monitoring and Control of Thinking and Reasoning". In: *Trends in Cognitive Sciences*.

Adleberg, T., M. Thompson, and E. Nahmias (2015). "Do Men and Women Have Different Philosophical Intuitions? Further Data". In: *Philosophical Psychology* 28.5, pp. 615–641.

Alexander, J and J. M. Weinberg (2014). "The "Unreliability" of Epistemic Intuitions". In: *Current Controversies in Experimental Philosohy*. Ed. by E Machery and E O'Neill. Routledge, pp. 128–145.

Alexander, J., C. Gonnerman, and J. Waterman (2015). "Salience and Epistemic Egocentrism: An Empirical Study". In: *Advances in Experimental Epistemology*. Ed. by J. Beebe. Continuum.

Alexander, J. et al. (2017). "Framing How We Think About Disagreement". In: *Philosophical Studies*, pp. 1–28.

Alicke, M., E. Gordon, and D. Rose (2013). "Hypocrisy: what counts?" In: *Philosophical Psychology* 26.5, pp. 673–701.

Anna-Sara, M. (2013). "Review of Herman Cappelen, Philosophy Without Intuitions". In: *Notre Dame Philosophical Reviews. http://ndpr. nd. edu/news/39362-philosophy-without-intuitions*.

Arango-Muñoz, S. and K. Michaelian (2014). "Epistemic Feelings, Epistemic Emotions: Review and Introduction to the Focus Section". In: *Philosophical Inquiries* 2.1, pp. 97–122.

Atlas, J. D. and S. C. Levinson (1981). "It-clefts, informativeness and logical form: Radical pragmatics (revised standard version)". In: *Radical pragmatics*. Academic Press, pp. 1–62.

Austin, J. L. (1962). *Sense and Sensibilia*. Oxford University Press.

Ayer, A. J. (1963). *The Foundations Of Empirical Knowledge*. Macmillan.

Ballantyne, N. and E. Coffman (2011). *Uniqueness, evidence, and rationality*. Ann Arbor, MI: Michigan Publishing, University of Michigan Library.

Baz, A. (2016). "Ordinary Language Philosophy". In: *The Oxford Handbook of Philosophical Methodology*. Ed. by H. Cappelen, T. S. Gendler, and J. Hawthorne. Oxford University Press Uk.

Bealer, G. (1998). "A Theory of Concepts and Concept Possession". In: *Philosophical Issues* 9, pp. 261+. DOI: 10\%2e2307/1522977. URL: http://dx.doi.org/10\%2e2307/1522977.

— (2000). "A theory of the a priori". In: *Pacific Philosophical Quarterly* 81, pp. 1–30.

Beebe, J. R. and W. Buckwalter (2010). "The Epistemic Side-Effect Effect". In: *Mind and Language* 25.4, pp. 474–498.

Bengson, J. (2013). "Experimental Attacks on Intuitions and Answers". In: *Philosophy and Phenomenological Research* 86.3, pp. 495–532.

— (2014). "How Philosophers Use Intuition and 'Intuition'". In: *Philosophical Studies* 171.3, pp. 555–576.

— (2015a). "A noetic theory of understanding and intuition as sense-maker". In: *Inquiry* 58.7-8, pp. 633–668.

— (2015b). "The Intellectual Given". In: *Mind* 124.495, pp. 707–760.

Blaauw, M. (2012). *Contrastivism in Philosophy: New Perspectives*. Routledge.

BonJour, L. (2013). "Is There a Priori Knowledge?" In: *Contemporary Debates in Epistemology*. Ed. by M. Steup and J. Turri. Blackwell, p. 177.

Boyd, K. and J. Nagel (2014). "The Reliability of Epistemic Intuitions". In: *Current Controversies in Experimental Philosophy*. Ed. by E. Machery and O. Elizabeth. Routledge, pp. 109–127.

Broad, C. D. (1923). *Scientific Thought*. Routledge and Kegan Paul.

Brown, J. (2006). "Contextualism and warranted assertibility manoeuvres". In: *Philosophical Studies* 130.3, pp. 407–435.

— (2013). "Cognitive Diversity and Epistemic Norms". In: *Philosophical Issues* 23.1, pp. 326–342.

— (2017). "The Gettier Case and Intuition". In: *Explaining Knowledge: New Essays on the Gettier Problem*. OUP.

Cappelen, H. (2012). *Philosophy Without Intuitions*. Oxford University Press.

— (2014). "Replies to Weatherson, Chalmers, Weinberg, and Bengson". In: *Philosophical Studies* .

Chalmers, D. J. (2014). "Intuitions in Philosophy: A Minimal Defense". In: *Philosophical Studies* 171.3, pp. 535–544.

Christensen, D. (2016). "Conciliation, Uniqueness and Rational Toxicity". In: *Noûs* 50.3, pp. 584–603.

Chudnoff, E. (2013). *Intuitions*. Oxford Univesity Press.

Cohen, S. (1999). "Contextualism, Skepticism, and the Structure of Reasons". In: *Philosophical Perspectives* 13.s13, pp. 57–89.

Colaço, D. and E. Machery (2017). "The intuitive is a red herring". In: *Inquiry* 60.4, pp. 403–419.

Colaço, D. et al. (2016). "On second thought: A refutation of the reflection defense". In: *Buffalo Experimental Philosophy Conference*.

Craig, E. (1990). *Knowledge and the State of Nature*. OUP.

Cummins, R. C. (1998). "Reflection on Reflective Equilibrium". In: *Rethinking Intuition*. Ed. by M. DePaul and W. Ramsey. Rowman & Littlefield, pp. 113–128.

Danek, A. H. and J. Wiley (2017). "What about false insights? Deconstructing the Aha! Experience along its multiple dimensions for correct and incorrect solutions separately". In: *Frontiers in psychology* 7, p. 2077.

Decker, J. (2012). "Disagreement, evidence, and agnosticism". In: *Synthese* 187.2, pp. 753–783.

Dennett, D. C. (1991). *Consciousness Explained*. Penguin Books.

DeRose, K. (1992). “Contextualism and Knowledge Attributions”. In: *Philosophy and Phenomenological Research* 52.4, pp. 913–929.

— (2011). “Contextualism, contrastivism, and X-Phi surveys”. In: *Philosophical studies* 156.1, p. 81.

Descartes, R. (1991). *The Philosophical Writings of Descartes, Vol. 3: Correspondence, Trans. By John G. Cottingham, Robert Stoothof, Dugald Murdoch, and Anthony Kenny.* Cambridge: Cambridge University Press.

Deutsch, M. (2009). “Experimental Philosophy and the Theory of Reference”. In: *Mind and Language* 24.4, pp. 445–466.

— (2015). *The Myth of the Intuitive*. The MIT Press.

Devitt, M. (2015). “Relying on Intuitions: Where Cappelen and Deutsch Go Wrong”. In: *Inquiry: An Interdisciplinary Journal of Philosophy* 58.7-8, pp. 669–699.

Douven, I. (2009). “Uniqueness revisited”. In: *American Philosophical Quarterly* 46.4, pp. 347–361.

Dretske, F. I. (1972). “Contrastive Statements”. In: *Philosophical Review* 81.4, pp. 411–437.

Earlenbaugh, J. and B. Molyneux (2009). “If Intuitions Must Be Evidential Then Philosophy is in Big Trouble”. In: *Studia Philosophica Estonica* 2.2, pp. 35–53.

Evans, J. S. B. T. and J Curtis-Holmes (2005). “Rapid responding inceases belief bias: Evidence for the dual-process theory of reasoning”. In: *Thinking & Reasoning* 11, pp. 382–389.

Evans, J. S. B., J. Clibbens, and B. Rood (1996). “The role of implicit and explicit negation in conditional reasoning bias”. In: *Journal of Memory and Language* 35.3, pp. 392–409.

Fantl, J. and M. McGrath (2002). “Evidence, Pragmatics, and Justification”. In: *Philosophical Review* 111.1, pp. 67–94.

Feltz, A. and E. Cokely (2012). “The Philosophical Personality Argument”. In: *Philosophical Studies* 161.2, pp. 227–246.

Ferretti, T. R., K. McRae, and A. Hatherell (2001). “Integrating verbs, situation schemas, and thematic role concepts”. In: *Journal of Memory and Language* 44.4, pp. 516–547.

Fischer, E. (2014). “Philosophical Intuitions , Heuristics , and Metaphors”. In: *Synthese* 191.3, pp. 569–606.

Fischer, E. and J. Collins, eds. (2015). *Experimental Philosophy, Rationalism, and Naturalism. Rethinking Philosophical Method*. Routledge.

Fischer, E. and P. E. Engelhardt (2016). “Intuitions’ Linguistic Sources: Stereotypes, Intuitions and Illusions”. In: *Mind and Language* 31.1, pp. 67–103.

Fischer, E. and P. E. Engelhardt (2017). “Stereotypical inferences: Philosophical relevance and psycholinguistic toolkit”. In: *Ratio* 30.4, pp. 411–442.

Fischer, E. and P. E. Engelhardt (forthcoming). “Stereotypical Inferences: Philosophical Relevance and Psycholinguistic Toolkit”. In: *Ratio*.

Fischer, E., P. E. Engelhardt, and A. Herbelot (2015). “Intuitions and Illusions: From Explanation and Experiment to Assessment”. In: *Experimental Philosophy, Rationalism, and Naturalism. Rethinking Philosophical Method*. Ed. by E. Fischer and J. Collins. Routledge, pp. 259–292.

Foley, R. (1992). *Working Without a Net: A Study of Egocentric Epistemology*. New York: Oxford University Press.

Freitas, J. D. et al. (2018). “Consistent Belief in a Good True Self in Misanthropes and Three Interdependent Cultures”. In: *Cognitive Science* 42.S1, pp. 134–160.

Gendler, T. (2010). *Intuition, Imagination, and Philosophical Methodology*. Oxford University Press.

Gerken, M. (2017). *On folk epistemology: How we think and talk about knowledge*. Oxford University Press.

Gerken, M. and J. R. Beebe (2016). “Knowledge in and Out of Contrast”. In: *Noûs* 50.1, pp. 133–164.

Gettier, E. (1963). “Is Justified True Belief Knowledge?” In: *Analysis* 23.6, pp. 121–123.

Ginzburg, J. (1995a). “Resolving questions, I”. In: *Linguistics and philosophy* 18.5, pp. 459–527.

— (1995b). “Resolving questions, II”. In: *Linguistics and Philosophy* 18.6, pp. 567–609.

Giora, R. (2003). *On our mind: Salience, context, and figurative language*. Oxford University Press.

Goldman, A. I. (2007). “Philosophical intuitions: Their target, their source, and their epistemic status”. In:

Greene, J. D. (2014). “Beyond Point-and-Shoot Morality: Why Cognitive (Neuro)Science Matters for Ethics”. In: *Ethics* 124.4, pp. 695–726.

Grimm, S. (2017). *Making Sense of the World: New Essays on the Philosophy of Understanding*. Oxford University Press.

Grimm, S. R. (2011). “Understanding”. In: *The Routledge companion to epistemology*. Routledge, pp. 110–120.

— (2014). “Understanding as knowledge of causes”. In: *Virtue epistemology naturalized*. Springer, pp. 329–345.

Grundmann, T. (2010). “Some Hope for Intuitions: A Reply to Weinberg”. In: *Philosophical Psychology* 23.4, pp. 481–509.

Hannon, M. (2013). “The practical origins of epistemic contextualism”. In: *Erkenntnis* 78.4, pp. 899–919.

— (2017). “A solution to knowledge’s threshold problem”. In: *Philosophical Studies* 174.3, pp. 607–629.

Hansen, N. (2014). “Contemporary Ordinary Language Philosophy”. In: *Philosophy Compass* 9.8, pp. 556–569.

— (2017). “Must We Measure What We Mean?” In: *Inquiry: An Interdisciplinary Journal of Philosophy* 60.8, pp. 785–815.

Hansen, N. and E. Chemla (2015). “Linguistic Experiments and Ordinary Language Philosophy”. In: *Ratio* 28.4, pp. 422–445.

Hare, M. et al. (2009). “Activating Event Knowledge”. In: *Cognition* 111.2, pp. 151–167.

Henderson, D. (2009). “Motivated contextualism”. In: *Philosophical Studies* 142.1, pp. 119–131.

— (2011). “Gate-keeping contextualism”. In: *Episteme* 8.1, pp. 83–98.

Henderson, D. K. and J. Greco (2015). *Epistemic Evaluation: Purposeful Epistemology*. OUP Oxford.

Hills, A. (2015). “Understanding Why”. In: *Noûs* 49.2, pp. 661–688.

Ichikawa, J. J. (2014). “Intuition in Contemporary Philosophy”. In: *Rational Intuition Philosophical Roots, Scientific Investigations*. Ed. by L. M. Osbeck and B. S. Held. Cambridge University Press.

— (2016). “Intuitive Evidence and Experimental Philosophy”. In: *Advances in Experimental Philosophy and Philosophical Methodology*. Ed. by J. Nado. Bloomsbury Academic.

Ichikawa, J. J. and B. Jarvis (2013). *Rules of Thought*. Oxford University Press.

Ichikawa, J. J. and M. Steup (2018). “The Analysis of Knowledge”. In: *The Stanford Encyclopedia of Philosophy*. Ed. by E. N. Zalta. Summer 2018. Metaphysics Research Lab, Stanford University.

Inwagen, P. van (1997a). “Materialism and the Psychological-Continuity Account of Personal Identity”. In: *Philosophical Perspectives* 11.s11, pp. 305–319.

Inwagen, P. V. (1997b). “Materiallism and the Psychological-Continuity Account of Personal Identity”. In: *Noûs* 31, pp. 305–319.

Jackson, F. (1982). “Epiphenomenal qualia”. In: *The Philosophical Quarterly (1950-)* 32.127, pp. 127–136.

— (2003). “Mind and illusion”. In: *Royal Institute of Philosophy Supplements* 53, pp. 251–271.

Kamm, F. M. (2015). *The trolley problem mysteries*. Oxford University Press.

Kauppinen, A. (2007). “The Rise and Fall of Experimental Philosophy”. In: *Philosophical Explorations* 10.2, pp. 95–118.

Kelly, T. (2011). “Following the argument where it leads”. In: *Philosophical Studies* 154.1, pp. 105–124.

— (2013). “Evidence can be permissive”. In: *Contemporary debates in epistemology*, pp. 298–312.

Kelp, C. (2011). “A problem for contrastivist accounts of knowledge”. In: *Philosophical studies* 152.2, pp. 287–292.

Kim, M. and Y. Yuan (2015). “No Cross-Cultural Differences in the Gettier Car Case Intuition: A Replication Study of Weinberg Et Al. 2001”. In: *Episteme* 12.3, pp. 355–361.

Knobe, J and S Nicholls, eds. (2008). *Experimental Philosophy*. Oxford: Oxford University Press.

Knobe, J. (2003). “Intentional action and side effects in ordinary language”. In: *Analysis* 63.279, pp. 190–194.

— (2016). “Experimental Philosophy is Cognitive Science”. In: *A Companion to Experimental Philosophy*. Ed. by J. Sytsma and W. Buckwalter. Blackwell.

Knobe, J. and Z. G. Szabó (2013). “Modals with a taste of the deontic”. In: *Semantics and Pragmatics* 6, pp. 1–1.

Koksvik, O. (2011). “Intuitions”. PhD thesis. Canberra, Australia: Australian National University.

Kopec, M. and M. Titelbaum (2016). “The Uniqueness Thesis”. In: *Philosophy Compass* 11.4, pp. 189–200.

Kripke, S. A. (1980). *Naming and Necessity*. Harvard University Press.

Kusch, M. and R. McKenna (2018). “The genealogical method in epistemology”. In: *Synthese*, pp. 1–20.

Leibniz, G (1973/1686). *Discourse on metaphysics*. Trans. by G Montgomery. La Salle, IL: Open Court Publishing.

Levin, I. P., S. K. Schnittjer, and S. L. Thee (1988). “Information framing effects in social and personal decisions”. In: *Journal of experimental social psychology* 24.6, pp. 520–529.

Levin, J. (2005). “The evidential status of philosophical intuition”. In: *Philosophical Studies* 121.3, pp. 193–224.

Lewis, D. (1983). *Philosophical Papers: Volume 1*. Oxford University Press.

Liao, S. M. et al. (2012). “Putting the Trolley in Order: Experimental Philosophy and the Loop Case”. In: *Philosophical Psychology* 25.5, pp. 661–671.

Littlejohn, C. (forthcoming). “A Plea for Epistemic Excuses”. In: *The New Evil Demon Problem*. Ed. by F. D. J. Dutant. Oxford University Press.

Lowe, E. J. (1983). “On the Identity of Artifacts”. In: *Journal of Philosophy* 80.4, pp. 220–232.

Ludwig, K. (2007). “The Epistemology of Thought Experiments : First Person Versus Third Person Approaches”. In: *Midwest Studies in Philosophy*. Ed. by P. A. French and H. K. Wettstein. Vol. 31. 1. Blackwell Pub. Inc., pp. 128–159.

Lycan, W. G. (1986). “Moral Facts and Moral Knowledge”. In: *Southern Journal of Philosophy* 24.S1, pp. 79–94.

Machery, E. (2015). “The Illusion of Expertise”. In: ed. by E. Fischer and J. Collins. Routledge.

Machery, E. (2017). *Philosophy within its proper bounds*. Oxford University Press.

Machery, E. et al. (2004). “Semantics, Cross-Cultural Style”. In: *Cognition* 92.3, pp. 1–12.

Markovits, H., V. A. Thompson, and J. Brisson (2015). “Metacognition and abstract reasoning”. In: *Memory & Cognition* 43.4, pp. 681–693. DOI: 10.3758/s13421-014-0488-9. URL: http://dx.doi.org/10.3758/s13421-014-0488-9.

Markovits, H. et al. (2017). “Interactions between inferential strategies and belief bias”. In: *Memory & cognition* 45.7, pp. 1182–1192.

Meacham, C. (2013). “Impermissive Bayesianism”. In: *Erkenntnis* S6, pp. 1–33.

Mizrahi, M. (2015). “Three Arguments Against the Expertise Defense”. In: *Metaphilosophy* 46.1, pp. 52–64.

Mogensen, A. L. (2016). “Contingency anxiety and the epistemology of disagreement”. In: *Pacific Philosophical Quarterly* 97.4, pp. 590–611.

Nado, J. (2014a). “Philosophical Expertise and Scientific Expertise”. In: *Philosophical Psychology* 28.7, pp. 1026–1044.

— (2014b). “Why intuition?” In: *Philosophy and Phenomenological Research* 89.1, pp. 15–41.

— (2015). “Intuition, philosophical theorizing, and the threat of skepticism”. In: *Experimental philosophy, rationalism, and naturalism: Rethinking philosophical method* 204.

— ed. (2016a). *Advances in Experimental Philosophy and Philosophical Methodology*. Bloomsbury Academic.

— (2016b). “Experimental Philosophy 2.0”. In: *Thought: A Journal of Philosophy* 5.3, pp. 159–168.

— (2016c). “The Intuition Deniers”. In: *Philosophical Studies* 173.3, pp. 781–800.

Nagel, J. (2010). “Knowledge Ascriptions and the Psychological Consequences of Thinking About Error”. In: *Philosophical Quarterly* 60.239, pp. 286–306.

— (2011). “The Psychological Basis of the Harman-Vogel Paradox”. In: *Philosopher’s Imprint* 11.5, pp. 1–28.

— (2012a). “Intuitions and Experiments: A Defense of the Case Method in Epistemology”. In: *Philosophy and Phenomenological Research* 85.3, pp. 495–527.

— (2012b). “Mindreading in Gettier Cases and Skeptical Pressure Cases”. In: *Knowledge Ascriptions*. Oxford University Press, pp. 171–191.

Nagel, J., V. San Juan, and R. A. Mar (2013). “Lay denial of knowledge for justified true beliefs”. In: *Cognition* 129.3, pp. 652–661.

Neely, J. H. (2012). “Semantic priming effects in visual word recognition: A selective review of current findings and theories”. In: *Basic processes in reading*. Routledge, pp. 272–344.

Newman, G. E., P. Bloom, and J. Knobe (2014). "Value judgments and the true self". In: *Personality and Social Psychology Bulletin* 40.2, pp. 203–216.

Neys, W. D. (2006). "Automatic-heuristic and executive-analytic processing during reasoning: Chronometric and dual-task considerations". In: *The Quarterly Journal of Experimental Psychology* 59.6, pp. 1070–1100.

Nichols, S. and J. Knobe (2007). "Moral Responsibility and Determinism: The Cognitive Science of Folk Intuitions". In: *Noûs* 41.4, pp. 663–685.

Nichols, S., S. Stich, and J. Weinberg (2003). "Metaskepticism: Meditations in ethno-epistemology". In: *The skeptics*, pp. 227–247.

Nozick, R. (1969). "Newcomb's Problem and Two Principles of Choice". In: *Essays in Honor of Carl G. Hempel*. Ed. by N. Rescher. Reidel, pp. 114–146.

Parfit, D. (1984). *Reasons and Persons*. Oxford University Press.

Picon, M. (2008). "What is the Foundation of Knowledge ? Leibniz and the Amphibology of Intuition". In: *Leibniz : What kind of Rationalist ?* Ed. by M. Dascal. Springer, pp. 213–227.

Pinillos, A. (2012). "Knowledge, Experiments, and Practical Interests". In: *Knowledge Ascriptions*. Ed. by J. Brown and M. Gerken. Oxford University Press, p. 192.

Pinillos, A. and S. Simpson (2014). "Experimental Evidence in Support of Anti-Intellectualism About Knowledge". In: *Advances in Experimental Epistemology*. Ed. by J. Beebe. Bloomsbury Academic, pp. 9–44.

Proust, J. (2013). *The Philosophy of Metacognition: Mental Agency and Self-Awareness*. OUP Oxford.

Pust, J. (2000). *Intuitions as Evidence*. Routledge.

Rini, R. A. (2015). "How Not to Test for Philosophical Expertise". In: *Synthese* 192.2, pp. 431–452.

Rose, D. and D. Danks (2013). "In defense of a broad conception of experimental philosophy". In: *Metaphilosophy* 44.4, pp. 512–532.

Rose, D., E. Machery, and S. S. et al (forthcoming). "Nothing at Stake in Knowledge". In: *Nous*.

Rosen, G. (2001). "Nominalism, Naturalism, Epistemic Relativism". In: *Nous* 35.s15, pp. 69–91.

Russell, B. (1912). *The Problems of Philosophy*. Barnes & Noble.

Rysiew, P. (2001). "The Context-Sensitivity of Knowledge Attributions". In: *Noûs* 35.4, pp. 477–514.

Scaltsas, T. (1980). "The Ship of Theseus". In: *Analysis* 40.3, pp. 152–157.

Schaffer, J. (2004). "From Contextualism to Contrastivism". In: *Philosophical Studies* 119.1-2, pp. 73–104.

— (2007). "Closure, contrast, and answer". In: *Philosophical Studies* 133.2, pp. 233–255.

Schaffer, J. and J. Knobe (2012). "Contrastive knowledge surveyed". In: *Noûs* 46.4, pp. 675–708.

Schoenfield, M. (2014). "Permission to believe: Why permissivism is true and what it tells us about irrelevant influences on belief". In: *Noûs* 48.2, pp. 193–218.

Schwarz, N., F. Strack, and H.-P. Mai (1991). "Assimilation and contrast effects in part-whole question sequences: A conversational logic analysis". In: *Public opinion quarterly* 55.1, pp. 3–23.

Schwitzgebel, E. and F. Cushman (2012). "Expertise in Moral Reasoning? Order Effects on Moral Judgment in Professional Philosophers and Non-Philosophers". In: *Mind and Language* 27.2, pp. 135–153.

Schwitzgebel, E. and F. Cushman (2015). “Philosophers’Biased Judgments Persist Despite Training, Expertise and Reflection”. In: *Cognition* 141, pp. 127–137.

Seyedsayamdost, H. (2015a). “On Gender and Philosophical Intuition: Failure of Replication and Other Negative Results”. In: *Philosophical Psychology* 28.5, pp. 642–673.

— (2015b). “On Normativity and Epistemic Intuitions: Failure of Replication”. In: *Episteme* 12.1, pp. 95–116.

Sinnott-Armstrong, W. (2011). “An empirical challenge to moral intuitionism”. In: *The new intuitionism*, pp. 11–28.

Sliwa, P. (2015). “IV—Understanding and Knowing”. In: *Proceedings of the Aristotelian Society* 115.1pt1, pp. 57–74.

— (2017). “Moral Understanding as Knowing Right from Wrong”. In: *Ethics* 127.3, pp. 521–552.

Smart, B. (1972). “How to Reidentify the Ship of Theseus”. In: *Analysis* 32.5, pp. 145–148.

Smith, A. D. (2002). *The Problem of Perception*. Harvard University Press.

Snedegar, J. (2017). *Contrastive Reasons*. Oxford University Press.

Sorensen, R. A. (1992). *Thought Experiments*. Oxford University Press.

Sosa, E. (2007a). “Experimental philosophy and philosophical intuition”. In: *Philosophical studies* 132.1, pp. 99–107.

— (2007b). “Intuitions: their nature and epistemic efficacy”. In: *Grazer Philosophische Studien* 74, pp. 51–67.

— (2009). “A defense of the use of intuitions in philosophy”. In: *Stich and his critics*, pp. 101–112.

Srinivasan, A. (2015). “The Archimedean Urge”. In: *Philosophical Perspectives* 29.1, pp. 325–362.

Sripada, C. S. and J. Stanley (2012). “Empirical Tests of Interest-Relative Invariantism”. In: *Episteme* 9.1, pp. 3–26.

Stahl, C., K. C. Klauer, and E. Erdfelder (2008). “Matching bias in the selection task is not eliminated by explicit negations”. In: *Thinking & Reasoning* 14.3, pp. 281–303.

Steglich-Petersen, A. (2015). “Knowing the answer to a loaded question”. In: *Theoria* 81.2, pp. 97–125.

Stich, S. (1990). *The Fragmentation of Reason: Preface to a Pragmatic Theory of Cognitive Evaluation*. MIT Press.

— (2009). “Replies”. In: *Stich and his critics*. Ed. by D. Murphy and M. Bishop. John Wiley & Sons.

Stich, S. and K. Tobia (2016). “Experimental Philosophy and the Philosophical Tradition”. In: *A Companion to Experimental Philosophy*. Wiley Blackwell, pp. 5–21.

Swain, S., J. Alexander, and J. Weinberg (2008). “The Instability of Philosophical Intuitions: Running Hot and Cold on Truetemp”. In: *Philosophy and Phenomenological Research* 76.1, pp. 138–155.

Sytsma, J. and W. Buckwalter (2016). *Blackwell Companion to Experimental Philosophy*. Wiley-Blackwell.

Thompson, V. A., J. A. P. Turner, and G Pennycook (2011). “Intuition, reason and metacognition”. In: *Conitive Psychology* 63, pp. 107–140.

— (2013). “The role of answer fluency and perceptual fluency as meta-cognitive cues for initiating analytic thinking”. In: *Cognition* 128, pp. 237–251.

Thompson, V. A., J. S. B. Evans, and J. I. Campbell (2013). "Matching bias on the selection task: It's fast and feels good". In: *Thinking & Reasoning* 19.3-4, pp. 431–452.

Thompson, V. A. and S. C. Johnson (2014). "Conflict, metacognition, and analytic thinking". In: *Thinking & Reasoning* 20.2, pp. 215–244.

Thomson, J. J. (1985). "The trolley problem". In: *The Yale Law Journal* 94.6, pp. 1395–1415.

Titelbaum, M. G. and M. Kopec (ms.). *When Rational Reasoners Reason Differently*.

Tobia, K., W. Buckwalter, and S. Stich (2013). "Moral Intuitions: Are Philosophers Experts?" In: *Philosophical Psychology* 26.5, pp. 629–638.

Tobia, K. P., W. Buckwalter, and S. Stich (2012). "Moral Intuitions: Are Philosophers Experts?" In: *Philosophical Psychology*, pp. 1–10.

Tsujii, T and S Watanabe (2010). "Neural correlates of belief-bias reasoning under time-pressure: A near-infrared spectroscopy study". In: *NeuroImage* 50, pp. 1320–1326.

Turri, J. and P. Blouw (2015). "Excuse validation: A study in rule-breaking". In: *Philosophical Studies* 172.3, pp. 615–634.

Van Fraassen, B. C. (1980). *The Scientific Image*. Oxford University Press.

Vavova, K. (2018). "Irrelevant Influences". In: *Philosophy and Phenomenological Research*, pp. 134–152.

Weatherson, B. (2014). "Centrality and Marginalisation". In: *Philosophical Studies* 171.3, pp. 517–533.

Weinberg, J. (2015). "Humans as instruments: Or, the inevitability of experimental philosophy". In: *Experimental Philosophy, Rationalism, and Naturalism. Rethinking Philosophical Method*. Ed. by E. Fischer and J. Collins. Routledge.

Weinberg, J. and J. Alexander (2014). "The Challenge of Sticking with Intuitions Through Thick and Thin". In: *Intuitions*. Ed. by A. R. Booth and D. P. Rowbottom. Oxford University Press.

Weinberg, J. M. (2007). "How to Challenge Intuitions Empirically Without Risking Skepticism". In: *Midwest Studies in Philosophy* 31.1, pp. 318–343.

— (2014). "Cappelen Between Rock and a Hard Place". In: *Philosophical Studies* 171.3, pp. 545–553.

Weinberg, J. M. (2016). "Intuitions". In: *The Oxford handbook of philosophical methodology*.

Weinberg, J. M. et al. (2010). "Are Philosophers Expert Intuiters?" In: *Philosophical Psychology* 23.3, pp. 331–355.

Weinberg, J. M. et al. (2012). "Restrictionism and Reflection: Challenge Deflected, or Simply Redirected?" In: *The Monist* 95.2, pp. 200–222.

White, R. (2010). "You Just Believe That Because. . . " In: *Philosophical Perspectives* 24.1, pp. 573–615.

Williamson, T. (2007). *The Philosophy of Philosophy*. Oxford: Wiley-Blackwell.

— (2011). "Philosophical Expertise and the Burden of Proof". In: *Metaphilosophy* 42.3, pp. 215–229.

— (2013). "Response to Cohen, Comesaña, Goodman, Nagel, and Weatherson on Gettier Cases in Epistemic Logic". In: *Inquiry: An Interdisciplinary Journal of Philosophy* 56.1, pp. 77–96.

Wright, J. (2010). "On Intuitional Stability: The Clear, the Strong, and the Paradigmatic". In: 115, pp. 491–503.

Zamzow, J. L. and S. Nichols (2009). "Variations in ethical intuitions". In: *Philosophical Issues* 19.1, pp. 368–388.