

Philosophy and Philosophy: The subject matter and the discipline

Ethan Landes



University of
St Andrews

This thesis is submitted in partial fulfilment for the degree of
Doctor of Philosophy (PhD)
at the University of St Andrews

December 2020

Abstract

The last two decades have seen the proliferation of the empirical study of philosophy. This dissertation defends the practice and argues that to understand the way contingent features of the practice of philosophy affect the epistemic standing of philosophers, we need to draw upon a wider and more varied set of empirical data than is sometimes supposed. To explore this, the dissertation focuses on two places where the practices of the discipline of philosophy have an effect on the epistemology of philosophy.

First, the dissertation discusses the interaction between notable works of philosophy and their readers. In particular, it critiques the method of defending the epistemic standing of philosophers through careful examination of notable works of philosophy to discern the methods in the text. Ultimately this method is epistemically unmotivated. It is instead far more important to study how people have interacted and reacted to works of philosophy.

Second, the dissertation defends the use of lexicography in philosophy. Using “intuition” as a case study, the dissertation argues metasemantically and lexicographically that philosophers often use common words with meanings unique to philosophy.

Through both discussions it is argued that experimental philosophers and epistemologists of philosophy need to drastically expand the sorts of data they collect and consider in their theorizing.

Chapter Abstracts

CHAPTER 1: INTRODUCTION

Chapter 1 introduces the background dialectic for the dissertation. Recent discussion on the epistemology of philosophy has centered around experimental studies of people's intuitions about thought experiments. While many experimental philosophers take their work to reveal the epistemic standing of philosophy, many philosophers disagree. Chapters 2, 3, and 4 in particular focus on the strategy developed by Max Deutsch and Herman Cappelen. Deutsch and Cappelen analyze presentations of thought experiments to argue that intuitions do not play an epistemic role in philosophy.

CHAPTER 2: THE THREAT OF THE INTUITION-SHAPED HOLE

Chapter 2 begins a critical discussion of Deutsch and Cappelen. They have not paid enough attention to how thought experiments have historically affected debates in philosophy. Careful exegesis of Gettier's cases reveals Gettier's argument (if there is one) is a bad argument. Combined with the historical features of post-1963 analytic epistemology, Deutsch and Cappelen are committed to widespread skepticism about philosophical beliefs. The chapter ends by generalizing this argument to other thought experiments.

CHAPTER 3: PRODUCER-FOCUSED METAPHILOSOPHY

Chapter 3 examines an assumption in Chapter 2, namely that examining original presentations of thought experiments reveals features of the epistemology of philosophy. Philosophical testimony would provide a defense of this method, as it creates a strong epistemic connection between the author of a work of philosophy and the reader of that work. However, the extent of peer disagreement within philosophy defeats any testimonial knowledge or justification we may gain from reading an original work of philosophy.

CHAPTER 4: HOW PHILOSOPHERS LEARN FROM EACH OTHER

Chapter 4 continues by examining what non-testimonial method philosophers use to learn from the works of other philosophers. Philosophers learn from each other through a process akin to pointing to arguments and reasons. The epistemic standing of readers of works of philosophy are therefore untethered from the *actual* epistemic standing of authors. Instead, readers' epistemic standing depends on what epistemic resources they themselves muster when reading works of philosophy. To develop our understanding of the epistemic standing of philosophy, we need to therefore develop tools for understanding the thought processes behind reading, not writing, philosophy.

CHAPTER 5: POLYSEMY

Chapter 5 develops a metasemantic account of when words have multiple related and established senses. Such polysemy, such as window (the glass opening) and window (the workspace on a computer screen), differs from homonymy such as bank/bank because the etymologies are related and the senses are connected in the minds of language users. Despite these connections, patterns of polysemy cannot be explained by mere differences in use, beliefs, or implicature. Instead, established polysemy should be treated like homonymy where different patterns of use represent different semantic meanings.

CHAPTER 6: JARGON

Chapter 6 explores the phenomena of technical language. Tying in discussion of the previous chapter, technical language often involves common English words used in polysemous ways. As with run-of-the-mill polysemy, polysemous senses only known to specialists also involve separate semantic meanings of a word. This has important consequences to methodological discussion of philosophy, as it provides a framework for understanding where and when philosophers use common English words with non-standard meanings.

CHAPTER 7: CASE STUDY: "INTUITION" AS POLYSEMY JARGON

Chapter 7 ties together all six substantive chapters of the dissertation. "Intuition" does not work like "water" or "arthritis" is thought to, where a single referent is discovered or determined by experts. Instead, experts – specifically philosophers working on intuitions – have aided in the process of splitting "intuition" into multiple related meanings. Drawing upon the framework developed in Chapters 5 and 6, this chapter argues this conclusion using lexicographic methods and criticizes extant lexicographic analyses of "intuition" by other philosophers.

Acknowledgements

Many people in my life have been invaluable mentors.

Thanks to my mother, whose hard work made this all possible. Thanks to my father, who gave me a healthy distrust of established thought. Thanks to John Dreher, my undergraduate advisor, who taught me the ways of crypto-pragmatism and instilled in me the conviction that philosophy is too broad and too valuable to spend one's day only reading philosophers. Thanks to Tom Ryckman, my undergraduate dissertation supervisor, who gave me my first tastes of doing real philosophy. Thanks to Marion Smiley for being a cynical voice of hope during my master's. Thanks to Jerry Samet, who, despite thinking my masters dissertation was "a tempest in a teapot", was willing to mentor me as I explored that tempest for the first time.

Thanks to Jessica Brown, my first PhD supervisor. Without Jessica the first half the dissertation would look nothing like it does now. Thanks to my longstanding second PhD supervisor Derek Ball. Derek has been a truly excellent supervisor the last four years. Thanks to Kevin Scharp, my current co-primary PhD supervisor. By introducing me to conceptual engineering, Kevin reminded me metaphilosophy can be fun. Thanks to Patrick Greenough, my other current co-primary supervisor. Patrick saw someone who had gotten lost in the woods of philosophy of language and needed a supervision shake-up. Without Patrick the second half of this dissertation would look nothing like it does now.

Thanks to Philip Ebert and Edouard Machery for a truly excellent viva.

Finally, thanks to my partner Sonja Dobroski and her daily reminders of how silly and powerful philosophy is. Sonja gives me hope for my future.

Many other people helped along the way. Thanks to James Andow, Lisa Bastian, Katharina Bernhart, Joseph Bowen, Rachel Cooper, Rasa Davidaviciute, Alex Douglas, Miguel Egler, Jeremy Fantl, Claire Field, Mirela Fus, Katherine Hawley, Manuel Gustavo Isaac, Bruno Jacinto, Tom Jones, Alison Duncan Kerr, Matt McGrath, Quentin Pharr, Mark Phelan, Lewis Ross, Fenner Tanswell, Deryn Thomas, Clotilde Torregrossa, Sara Vikesdal and all of the participants at the many, many seminars at Arché and SASP where I presented early versions of this work.

Candidate's declaration

I, Ethan Landes, do hereby certify that this thesis, submitted for the degree of PhD, which is approximately 73,000 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 confirm that any appendices included in my thesis contain only material permitted by the 'Assessment of Postgraduate Research Students' policy.

I was admitted as a research student at the University of St Andrews in September 2016.

I confirm that no funding was received for this work.

Date: 11 May 2021

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. I confirm that any appendices included in the thesis contain only material permitted by the 'Assessment of Postgraduate Research Students' policy.

Date

17 May 2021

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, Ethan Landes, 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 of this material is finished.

Title and Abstract

- I agree to the title and abstract being published.

Date: 11 May 2021

Signature of candidate:



Date 17 May 2021

Signature of supervisor



Underpinning Research Data or Digital Outputs

Candidate's declaration

I, Ethan Landes, 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: 11 May 2021

Signature of candidate: 

Contents

1	Introduction	1
1.1	Introducing the Negative Thesis	4
1.2	Experimental Philosophy	5
1.3	The Experimental Attack	12
1.4	Intuition Denial	18
1.5	Conclusion	26
2	The Threat of the Intuition-Shaped Hole	28
2.1	The Threat of the Intuition-Shaped Hole	30
2.2	The 10 Coins Case	32
2.3	Defending the Verdict Without Intuitions	34
2.4	Abduction in Gettier	42
2.5	Other Thought Experiments	49
2.6	Conclusion	56
3	Producer-Focused Metaphilosophy	59
3.1	Introduction: Exegesis vs Xphi	60
3.2	Focusing on Producers	66
3.3	Existing Defenses	69
3.4	Testimony and Philosophy	73
3.5	Peer Disagreement and Philosophical Testimony	81
3.6	Conclusion	86
4	How Philosophers Learn From Each Other	88
4.1	Returning to Producers and Consumers	89
4.2	Showing	91
4.3	Exegesis Reveals Next to Nothing	94
4.4	Fake Philosophy is Real Philosophy	98
4.5	Objection: “Showing” is Just Testimony	105
4.6	The Importance of Empirical Facts	110
4.7	Conclusion	114

CONTENTS

5	Polysemy	116
5.1	Introducing Polysemy	117
5.2	Introducing Meaning Polysemy	122
5.3	Master Argument: Polysemy is not special	126
5.4	Use Polysemy	137
5.5	Belief Polysemy	138
5.6	Implicature Polysemy	142
5.7	Conclusion	146
6	Jargon	148
6.1	Introducing Jargon	149
6.2	Existing Approaches to Identifying Jargon	153
6.3	Polysemy Jargon	162
6.4	Conclusion	179
7	Case Study: “Intuition” As Polysemy Jargon	180
7.1	Introduction	180
7.2	Previous Metaphilosophical Analyses of “Intuition”	182
7.3	“Intuition” as Polysemy Jargon	194
7.4	Objection 1: Return of the Semantic Anarchy Worry	199
7.5	Objection 2: Metalinguistic Claims in Bengson (2014)	205
7.6	Moving the Intuition Debate Forward	209
7.7	Conclusion: Beyond One Meaning	216
8	Final Thoughts	219
8.1	Future Directions	220
8.2	The End	222

Chapter 1

Introduction

As the title suggests, this dissertation is about the interaction between the subject matter of philosophy and the academic discipline of philosophy. Philosophy is ultimately a human enterprise. As much as we might wish we were truth-seeking philosopher kings or Cartesian ghosts unburdened by life, our philosophizing is instead the product of psychology, culture, disciplinary norms, and other contingent features of ourselves and our environment. This fact is often lost in how philosophy is done and, as is discussed here, how philosophers theorize about their own practice.

Analytic philosophers have long seen the practice of reconciling intuitions with general principles as central to philosophy. This has been viewed as an essentially armchair practice that needs nothing more than a philosopher and a well thought-out thought experiment. The turn of the millennium saw a rejection of this view of philosophy by experimental philosophers, who argued that the question of what is intuitive is empirically testable. Some experimental philosophers went a step further and argued that the empirical data they collected on intuitions showed that philosophers' intuitions about thought experiments were bad evidence for the truth of general principles.

This understandably did not sit well with many philosophers. With an eye towards defusing the skeptical worries raised by such experimental philosophers, a number of

critics made the surprising move of denying the role of intuitions in philosophy. Such intuition deniers argued that philosophers have been mistaken in talking about intuitions, since philosophers do not and have not used intuitions as evidence for philosophical claims.

This dissertation explores the relationship between the subject matter of philosophy and the discipline of philosophy by defending two main theses about this debate. The first thesis is negative. I defend experimental philosophy against a version of intuition denial defended by Deutsch (2015) and Cappelen (2012). Deutsch and Cappelen reject intuitions' role of philosophy through analysis of how philosophers have presented influential thought experiments in print. I argue that this strategy is fundamentally broken. For one, it fails to defuse epistemic worries raised by experimental philosophy because the strategy is itself committed to widespread skepticism about how philosophers have come to believe the principles they believe. More problematically, there is no epistemic motivation for examining original texts in the first place. Readers of philosophy do not gain justification based on what a work *actually* says. Rather they gain justification based on their own reading and analysis of the text.

My positive thesis builds upon this by arguing that there is something nonetheless right about the spirit of Deutsch and Cappelen's approach. Philosophers interested in the epistemology of philosophy need to draw from a wide array of contingent facts about how philosophy is done. Therefore, I argue, philosophers interested in the methods and epistemology of philosophy need to draw upon fields like sociology, anthropology, and lexicography.

Chapters 2, 3, and 4 are primarily focused on the negative thesis, although I do forward a number of positive claims about the sorts of contingent facts that matter to the epistemology of philosophy. In those chapters I develop two independent lines of criticism against Deutsch (2015) and Cappelen (2012) and their methods. I argue first in Chapter 2 that by denying intuition's role in philosophy in the way they do, they

are committed to widespread skepticism about the philosophical beliefs contemporary philosophers hold.

Then in Chapter 3, I begin exploration of whether the strategy of examining original texts discussed in Chapter 2 is motivated in the first place. I argue in Chapter 3 that philosophers do not learn from each other via testimony – which is bad news for Deutsch and Cappelen. In Chapter 4, I then argue communication between philosophers is more akin to pointing or showing, rather than telling, meaning the epistemically important facts are about how works of philosophy are understood by *readers* or *listeners* rather than *authors*.

After those three chapters, the dissertation’s focus shifts to the philosophy of language of philosophy and defending the positive claim that lexicography is important to the study of philosophy’s methods. I ultimately argue that “intuition” should not be taken to have one meaning in the mouths of epistemologists of philosophy, but my claim needs significant set-up.

Chapter 5 develops a metasemantic story for polysemy, which is when a word has multiple related but distinguishable senses. I argue that these different senses correspond to different literal meanings of a word. Chapter 6 then applies this account to technical language. Drawing upon examples in philosophy and other technical areas, I argue that technical language often involves co-opting common English words and giving them additional meanings. Therefore, common words often have literal meanings that are only known by people with high levels of technical knowledge. With this framework in place, I examine “intuition” in Chapter 7. I argue that “intuition” has multiple technical and literal meanings, or at least multiple candidates for new and additional meanings. I end by arguing that the study of the epistemology of philosophy should acknowledge and respect this.

1.1 Introducing the Negative Thesis

In the rest of this chapter, I introduce the background to the negative thesis. In particular, I set up the debate between certain strands of EXPERIMENTAL PHILOSOPHY and what I call EXEGESIS-BASED INTUITION DENIAL defended by Deutsch (2015) and Cappelen (2012).¹ The former argues experimental data on how people react to thought experiments can tell us about the epistemic standing of philosophy. Such experimental philosophers think that empirical evidence that people's intuitions have certain features is evidence that intuition's use in philosophy is epistemically problematic. Exegesis-based intuition deniers, in contrast, reject this line of reasoning put forth by certain experimental philosophers and their allies. Intuition deniers instead argue that if we look carefully at how philosophers have presented thought experiments, intuitions were not used as a source of evidence. Instead, they claim, authors *argue* for the thought experiment's verdict. Therefore, the deniers argue, experimental data about reactions to thought experiments cannot tell philosophers anything about the epistemic standing of other philosophers.

The plan for this introductory chapter is as follows. In Section 1.2, I introduce the project of experimental philosophy, moving to specifically focus on what has classically been called negative experimental philosophy. In Section 1.3 I discuss how it is that experimental data on people's reactions to thought experiments are thought to put pressure on the epistemic standing of philosophy. This is a four-part argument I call the EXPERIMENTAL ATTACK. From there, in Section 1.4 I introduce intuition denial, the strategy of defusing the experimental attack by rejecting the experimental attack's premise that, as a matter of fact, intuitions play a role in philosophy. While intuition denial has multiple forms, I specifically introduce the method of using exegetical analysis of original presentations of thought experiments in order to reject intuition's role in philosophy.

¹ See also Cappelen (2014a, 2014b); Deutsch (2009, 2010, 2016).

This introduction lays the groundwork for two distinct lines of attack on exegesis-based intuition denial. I argue in Chapter 2 that while intuition denial is meant to protect philosophy from skeptical worries raised by experimental philosophy, it is itself committed to nearly identical skeptical worries. Classic presentations of thought experiments successfully changed philosophers' minds, and Deutsch and Cappelen claim (and are committed to the claim) that those texts contain bad arguments. This then means that they are committed to the claim that philosophers formed their philosophical beliefs based on bad arguments, which is potentially worse than the claim they tried to avoid – that philosophers formed their beliefs based on bad intuitions.

In Chapters 3 and 4, I then argue against the metaphilosophical relevance of looking at original texts in the first place. While it might be interesting from a historical perspective to understand how Gettier or Thomson used and understood thought experiments, it is not in itself relevant for the purposes of evaluating the epistemic standing of philosophy and philosophers. Based on how it is philosophers learn from each other, which is *not* via testimony, I argue focus should instead be on what epistemic attitude experts who read the original texts take towards the texts rather than on what the texts themselves actually say.

1.2 Experimental Philosophy

To understand intuition denial, we need to understand the movement it was developed in response to, namely, experimental philosophy. Experimental philosophy is a loosely connected and two-decades-old tradition of philosophy united by the use of empirical methods on topics and debates traditionally thought to belong to the realm of armchair reflection. Instead of approaching philosophy through exclusively armchair-based reflection and argumentation, experimental philosophers have instead borrowed or built upon a range of empirical methods from other fields, especially psychology.

Traditional experimental philosophy studies follow a basic methodological structure. Experimental philosophers choose an existing debate or topic within philosophy they would like to empirically investigate. Their goal is to either tweak how a thought experiment is presented, what question is asked about the thought experiment, or who responds to the thought experiment in order to learn something about the debate or topic. If the debate has suitable thought experiments already written by philosophers, experimental philosophers borrow these thought experiments. This can either be verbatim or tweaked slightly to aid participants' comprehension. If existing thought experiments are not suitable (e.g., they are too complicated, they do not capture what the experimenter is interested in, or they have been tested to death), the experimenter develops their own. The experimenter then distributes the thought experiments to a large audience, aggregating and analyzing the results.

Turri (2016) is a fairly typical example of the form. Turri is interested in experimentally examining the knowledge norm of assertion and the certainty norm of assertion. These are the claims that we should only assert something if we know or are certain of it. To draw the knowledge norm of assertion and the certainty norm of assertion apart, Turri develops his own two sets of thought experiments. One, called CABIN, is the following:

Angelo is camping with his daughter in a wooden cabin at the edge of the forest. As they settle in to sleep for the night, the daughter has her headphones on and Angelo is reading near the window. Angelo hears two very loud, sharp bangs ring out in the forest behind the cabin. It is deer-hunting season. Angelo's daughter takes off her headphones and asks, "Dad, what's going on? Is somebody hunting deer nearby?"

Either Cabin or a related case about a human resources office was given to 182 participants. Participants were then shown four statements about Angelo and asked to select which one is true. One set of participants saw statements about whether Angelo knows

and another saw statements about whether Angelo is certain:

1. He [knows/is certain] that someone is hunting nearby, and he should say that someone is hunting nearby.
2. He [knows/is certain] that someone is hunting nearby, and he should not say that someone is hunting nearby.
3. He [does not know/is not certain] that someone is hunting nearby, and he should say that someone is hunting nearby.
4. He [does not know/is not certain] that someone is hunting nearby, and he should not say that someone is hunting nearby.

Turri analyzes the experiment's results by comparing whether or not participants tend to connect knowing and assertion (i.e., they chose the first or last option) or think they can come apart (i.e., they chose the second or third option). Turri then compares the differences in responses between participants who saw statements about certainty and those who saw statements about assertion. Analysis reveals that participants are more likely to answer in ways consistent with the knowledge norm of assertion than the certainty norm of assertion, which Turri takes as support for the knowledge norm.

Case-based experiments such as Turri (2016) that analyze answers about thought experiments are paradigmatic examples of experimental philosophy, and they will be the focus of discussion here. Nonetheless it is worth pointing out that recent years have seen an explosion of new experimental methods. Some experimental projects have worked to go beyond the survey, running experiments in situ. Meskin and Liao (2018) test the relevance of tasting notes in aesthetic judgements by hosting a coffee tasting at a coffee shop. Similarly, Schwitzgebel (2009) tests the morality of professional ethicists by observationally testing the rates of stolen library books about ethics compared to other disciplines (see also Rust & Schwitzgebel, 2013; Schwitzgebel & Rust, 2014). Other projects have

dropped surveys altogether. Using corpora to compare how philosophers talk against how they theorize, Hansen, Porter, and Francis (2019) examine philosophers' use of "know" and Andow (2015a); Ashton and Mizrahi (2018) examine philosopher's use of intuition-talk. Even surveys built around thought experiments have seen new methods in recent years. Fischer and Engelhardt (2017b) use pupillometry (measurements of the size of participants' pupils) to examine the mental effort behind responses to surveys. Hannikainen, Machery, and Cushman (2018) uses the Moral Sense Test, a free online moral inventory built around short descriptions of hypothetical examples, to test whether people's moral judgements changed over roughly 10 years. The maturation of experimental philosophy has also enabled the production of qualitative meta-analyses aggregating and analyzing dozens of surveys (Machery, 2017). This includes Cova et al. (2018), who, by attempting to replicate a battery of previous experimental philosophy, ran a meta-analysis of experimental philosophy's replicability.

Writing at the start of this proliferation of methods, Weinberg (Weinberg, 2015, 172-173 fn3) attempts to tie these together under the tentative definition of experimental philosophy as the practice of "recognizing philosophy's empirical commitments, wherever they may be found, and applying the best methods available for evaluating such commitments, whatever they may be". In here is a strong rebuttal of a traditional – but by no means ever universal – self-conception of analytic philosophy as philosophy that can be done with only half an eye towards empirical claims.² In particular, early experimental philosophers took issue with the apparent supposition in arguments containing thought experiments that what is intuitive to well-off Anglo-American philosophers is intuitive to everyone. That project, while still ongoing (e.g., Barrett, 2020a, 2020b), is nonetheless a small part of a much larger tradition. Experimental philosophers continue to find ways to hunt down purportedly empirical commitments in *apparently* armchair-evaluable

² Bealer (1998) is a classic of the form. See Fuller (2002) for a discussion of the driving forces behind this self-conception.

arguments and test those commitments themselves.

Weinberg's definition is not perfect for at least three reasons. First, as Weinberg himself admits, his definition is too exclusive as it is now common for experimental philosophers to run studies that are merely meant to forward our understanding of a notion as opposed to trying to evaluate some pre-existing philosophical claim. Second, it is not clear we can give necessary and sufficient conditions for experimental philosophy. Experimental philosophy is now a tradition of philosophy in its own right, much like pragmatism or ordinary language philosophy, and like those traditions, practitioners of experimental philosophy have competing methodologies and self-conceptions. Trying to unite all the different strands of the tradition in any informative and counterexample-proof way may therefore prove impossible. Third, as experimental philosophy has progressed and interdisciplinary collaborative teams have become more common, the line between works of experimental philosophy and, say, works of corpus linguistics have become increasingly blurred.

An extensionally adequate definition would, for the three reasons just discussed, likely be nothing more than *empirical projects led by people who identify as philosophers in their capacity as philosophers*. Nonetheless, Weinberg's definition is instructive even if it only captures the spirit, if not every instance, of experimental philosophy. As a tradition, experimental philosophy spots and tests philosophy's implicit or explicit empirical claims by using experimental tools. While the claims and tools were originally limited to psychology, the tradition has become more methodologically diverse as the tradition has matured.

In this way, this dissertation is embedded within the tradition of experimental philosophy. I indicate new directions experimental philosophers should explore. For one, they should focus on the interactions between philosophers. Experimental philosophers should devise new methods to test the ways philosophers read and understand works of philos-

ophy that go beyond just testing the reactions philosophers have when they read thought experiments. We need to study the epistemic attitudes philosophers take towards texts – whether they trust and take on what philosophers assert as testimony – and study historical evidence about how specific texts have influenced debates. Experimental philosophers should also develop methods, likely in corpus linguistics, that spot when philosophers use words with meanings unique to philosophy.

1.2.1 Negative Experimental Philosophy

Discussion on how to broaden empirical methods will come later. For now, my focus is on introducing the intuition denial at the heart of this dissertation’s negative thesis. Intuition denial is not interested in the whole of experimental philosophy or even the sort of study that Turri (2016) exemplifies. Rather, intuition denial is primarily interested in experimental philosophy that purports to show something about the epistemic standing of philosophers’ beliefs and the methods they use to arrive at those beliefs.

When I started writing this dissertation, it still was common for people to distinguish between positive and negative experimental philosophy. Positive experimental philosophy (such as Turri (2016)) was thought to be in service of traditional case-based armchair philosophy. This was contrasted with negative experimental philosophy, which was seen as trying to burn armchair philosophy to the ground. The label “negative experimental philosophy” was for many years apt and well-deserved. Negative experimental philosophy’s splashiest and earliest works were decidedly antagonistic towards the use of traditional armchair methods.

In recent years, however, the name “negative experimental philosophy” has fallen out of favor as much of the work done in the tradition of negative experimental philosophy has cooled its polemical stance towards armchair philosophy. Most people working in the tradition today see themselves as allies of armchair philosophy, even if they do

not endorse unrestricted use of thought experiments. Current research in the negative tradition is now largely focused on mapping out the pitfalls of intuitions and adjusting practice accordingly, although there is substantive disagreement about how widespread those pitfalls are. Accordingly, the label “negative experimental philosophy” fails to capture the current spirit of the tradition. I instead prefer to call this tradition “evaluative experimental philosophy” to capture the distinctly normative dimension of the work (for recent examples of work in the evaluative tradition, see Fischer, Engelhardt, Horvath, and Ohtani (in press); Machery (2017); Mortensen and Nagel (2016)).³

One strand of the evaluative tradition has looked at how contingent features of oneself – such as culture or personality – affect responses to thought experiments. Some of the earliest and splashiest findings in the tradition found intuitions differed by socioeconomic status (Weinberg, Nichols, & Stich, 2001) and western vs eastern culture (Machery, Mallon, Nichols, & Stich, 2004; Weinberg et al., 2001). Later studies in this tradition also found effects driven by gender (Buckwalter & Stich, 2010) and personality (Bartels & Pizarro, 2011; Schultz, Cokely, & Feltz, 2011). This strand has hit some setbacks in recent years. In many of these early studies, sample sizes were small, participants were all undergraduates at US universities, and statistical analyses were problematic.⁴ As experimental philosophy has matured, so has its methods, including sampling practices, statistical analysis, and the adoption of pre-registration of experimental design (e.g., Cova et al. (2018)). When these findings were tested again, some findings did not replicate, especially those involving gender and Gettier cases (Adeberg, Thompson, & Nahmias, 2015; Nagel, San Juan, & Mar, 2013; Seyedsayamdost, 2015a, 2015b). Nonetheless there is still significant evidence that culture and personality affect intuitions about thought experiments other than Gettier cases (Machery, 2017, 48-69).

³ An early version of this was called “restrictionism” (Alexander & Weinberg, 2007; Weinberg, 2007). I do not like this name either given the possibility of using experimental philosophy to improve our intuitions as opposed to cutting the bad intuitions out (see Andow, 2018).

⁴ For discussion of this last point, see Ebert, Smith, and Durbach (2018).

A different strand of the evaluative tradition has examined whether *how* cases are presented impacts intuitions. This strand has found that cases that are structurally identical from a philosophical point of view do not necessarily result in identical intuitions. While results differ quite a bit by subfields within philosophy, studies have found some intuitions are affected by the order in which cases are presented (e.g., Cokely & Feltz, 2009; Liao, Wiegmann, Alexander, & Vong, 2012; Swain, Alexander, & Weinberg, 2008),⁵ features of the environment in which cases are considered (e.g., Helzer & Pizarro, 2011; Tobia, Buckwalter, & Stich, 2013), and small, supposedly irrelevant tweaks to the language of the cases (e.g., Machery et al., 2018; Nadelhoffer & Feltz, 2008; Uhlmann, Pizarro, Tanenbaum, & Ditto, 2009).

1.3 The Experimental Attack

The evaluative experimental project uses empirical data to infer the epistemic status of certain intuitions. Particularly, the evaluative experimental project claims that certain intuitions are ill-suited for philosophical theorizing. Call the multi-step argument from empirical data to a worry about the epistemic standing of philosophy the EXPERIMENTAL ATTACK. The experimental attack uses data about the psychology of philosophical activity to raise second-order worries about the epistemic status of intuitions (Goldman, 2010). Proponents of the experimental attack do not think that surveys of peoples' intuitions about knowledge tell us what knowledge is. Rather, evaluative experimental philosophers take experimental data to tell us whether or not intuitions about knowledge are good evidence for our theories about knowledge.

To see this in action, consider the pair of studies, Schwitzgebel and Cushman (2012,

⁵ To add extra complexity, order effect appears to depend, at least sometimes, on whether people are prompted to ignore order (Cullen, 2010). This cannot be the whole story, however, as professional philosophers – who presumably know that order does not matter – are also affected by order (Schwitzgebel & Cushman, 2015).

2015). Schwitzgebel and Cushman start with three variations of trolley cases: one where a switch is toggled to divert the trolley (SWITCH), one where a person is pushed into the path of the trolley (PUSH), and one where a switch is toggled to drop someone into the path of the trolley (DROP). Participants saw Switch alongside either Push or Drop, and the two cases appeared in a random, but tracked, order. One of the main findings of both the 2012 and 2015 studies is that whether participants judged flipping the switch in Switch and pushing the person in Push as morally equivalent depended on the order in which the two cases were presented. When Switch was presented before Push (which is, notably, the order Thomson (1976b) presents them) participants were less likely to view the sacrifice as morally equivalent than if the order was flipped and Push was presented before Switch. Importantly, Schwitzgebel and Cushman (2012, 2015) tested both philosophers and non-philosophers, finding the effect in both groups.⁶

Schwitzgebel and Cushman end their 2012 paper by saying:

The method of philosophy is often characterized as a matter of reconciling intuitive judgements about particular cases with plausible general principles [...] Our results suggest that even professional philosophers' judgements about familiar types of cases in the their own field can be strongly and covertly influenced by psychological factors that they would not endorse upon reflection [...] (2012, 150) ⁷

In this quote we can find a relatively clear articulation of the experimental attack:

1. Philosophy involves reconciling intuitions with principles
2. According to our findings, these intuitions are affected by order effects

⁶ See Devitt (2011); Ludwig (2007); Rini (2015); Williamson (2007) for the motivation behind such studies. See also Egler and Ross (2020); Nado (2014, 2015); Weinberg, Gonnerman, Buckner, and Alexander (2010).

⁷ I have omitted their discussion of the findings that order effects later also influenced what moral principles philosophers later endorsed. This particular finding did not replicate in Schwitzgebel and Cushman (2015).

3. Philosophers would not endorse order effects upon reflection
4. [Implicit] Therefore, philosophy is in bad epistemic standing

Premise 3 is a specific version of a more general move, in which certain features of intuitions are inferred to make them ill-suited for the work philosophers place on them. Schwitzgebel and Cushman’s language of “endorsing under reflection” has the virtue of being theoretically light-weight since it makes no assumption about whether there is a fact of the matter about whether the effects are problematic. Rather than appealing to facts of the matter about the moral status of actions in Switch and Push, they merely point out that by philosophers’ own lights the effect is bad.

While I am sympathetic to this sort of presentation of ill-suitedness for reasons I articulate in Section 1.3.1, it does not match the language of the other authors I am considering here. In the metaphilosophical literature, ill-suitedness has cashed out in terms of being sensitive to factors irrelevant to truth (Alexander & Weinberg, 2007; Deutsch, 2015; Fischer & Engelhardt, 2017a; Horvath, 2010), being unreliable (Boyd & Nagel, 2014; Machery, 2017), being hopeless (Weinberg, 2007), being unrepresentative (Cappelen, 2012), among others. Since the next few chapters focus specifically on Deutsch’s formulation of exegesis-based intuition denial, I will follow him and talk about truth-irrelevance.

Reordering the experimental attack slightly, adjusting it to be about truth-irrelevance, and broadening discussion of problematic effects to include things other than order effect, we end up with:

1. According to experimental findings, intuitions are affected by P
2. P is truth-irrelevant
3. Philosophy involves reconciling intuitions with principles
4. Therefore, philosophy is in bad epistemic standing

To be valid, this presentation of the experimental attack needs an additional principle or two bridging the premises and the conclusion. Presenting a deductively valid version of the experimental attack is not my goal here, however. Rather, I have formulated the experimental attack this way to highlight the three steps with three very different types of claims needed to establish the skeptical conclusion in 4. The first step is a claim about what *psychological data* experimental philosophers have found. The second step is a claim about how *philosophical truth* relates to the findings in the first step. The third step is a claim about what *methods* philosophers actually use. I have already discussed the sorts of empirical data used in step 1, and step 3 will be the focus of the next three chapters. It is therefore worth briefly looking at how step 2 is established.

1.3.1 Truth-Irrelevance

The experimental attack owes its origin to the heuristics and biases literature, which has found a wide range of problematic sensitivities in domains from statistics to sentencing by judges to decisions of hiring committees (for introduction, see Kahneman, 2013). Imagine we want to judge whether people’s statistical judgements are good judgements. In the classic example of Linda the bank teller (Tversky & Kahneman, 1983), participants are told that 31 year-old Linda does “alternative” things, such caring about social justice, caring about discrimination, and taking part in anti-nuclear demonstrations. Linda is also described as being intelligent, outspoken, single, and holding a bachelor’s degree in philosophy. Participants are then asked which of the following is more likely:

A: Linda is a bank teller

B: Linda is a bank teller and is active in the feminist movement

In the original study, 85 percent of participants said that B is more likely than A. However, given classical statistics, the probability of a conjunction is less than or equal to

the probability of the conjuncts. Assuming no problems with the experimental design,⁸ these findings allow the fairly straightforward inference that people are bad at these sorts of statistical inferences. According to statistics, P is true, but people judge not-P. Therefore people should be careful about their judgements in Linda-like cases. Put in terms of truth-irrelevance, the truth-relevant feature of the question or puzzle participants need to consider is the conjunction of two less-than-certain possibilities. Instead, judgements track the perceived higher likelihood that Linda is a feminist rather than being a bank teller, something that is irrelevant to the truth of which is more likely.

While this method is the basis of experimental philosophy's experimental attack, there is an important disanalogy between the two. Classical statistics is an agreed-upon standard of truth for evaluating statistical judgments, but philosophers lack the same external arbiter of truth for intuitions about cases. Besides propositional logic, which is of little help in determining the truth of thought experiments, we do not have an agreed-upon objective standard by which to judge intuitions. The experimental attack instead relies on the assumption that non-experimental philosophical methods are somewhat truth-tracking already.

This is not a robust claim of truth-tracking where the experimental attack assumes certain philosophical theories are right. Philosophers might not all have correct beliefs about the facts of the matter, but the experimental attack assumes that they roughly know what sort of theory is true and what sort of theory is false. Since theories are developed to explain certain features of thought experiments (Andow, 2016a; Williamson, 2016), for the experimental attack to work, philosophers therefore need to have some correct idea about what features of a thought experiment are truth-relevant.

To return to trolley cases, ethicists do not agree about what the morally permissible

⁸ I have serious concerns about the design as stated. I think it is likely that pragmatic mechanisms cause participants to infer the content of [A] is that "Linda is a bank teller and *not* a feminist". For an introduction to pragmatic worries in experimental design, see Cullen (2010) and Deutsch (Deutsch, 2015, 21-24). I am setting these worries aside for purposes of exposition. Thanks to Rasa Davidaviciute and Deryn Thomas for discussion on the point.

action in certain variations are. Nor do they agree whether we should theorize about them from, say, a rule-utilitarian or Kantian perspective. Nonetheless, philosophers generally agree that something like rule-utilitarianism or Kantian deontology is correct. They also generally agree about which features of the cases are plausibly truth-relevant. Plausibly truth-relevant features of trolley cases include the numerical trade-off of lives and the intention of the person making the decision. When intuitions about thought experiments are found to be sensitive to features that are not plausibly truth-relevant, then intuitions are said to track truth-irrelevant features. Returning to order effects, the order in which cases are presented is not taken to be a plausibly truth-relevant feature so it is taken to be problematic that intuitions are affected by it.

If armchair practices lead to different conclusions about what is plausibly truth-relevant, then different sensitivities would motivate the experimental attack. To illustrate this point, consider another example of findings in intuitions about trolley cases. Uhlmann et al. (2009) found that trolley case judgments differ depending on race. Participants were either asked whether it is justified to push Chip Elseworth III (a stereotypically white name) to his death to save 100 members of the Harlem Jazz Orchestra (a stereotypically black ensemble) or whether it is justified to push Tyrone Payton (a stereotypically black name) to his death to save 100 members of the New York Philharmonic (a stereotypically white ensemble). Uhlmann et al. (2009) found that race affected participants' choices, but how it affected participants' choices depended on the participants' political orientation. Liberal-leaning participants were more likely to sacrifice Chip Elseworth III and conservative-leaning participants were more likely to sacrifice Tyrone Payton. The effect was found despite participants reporting that race should not affect decisions like this.

To simplify discussion, remove the role of political orientation and imagine a hypothetical "Study R" that found that across the board people were more likely to sacrifice the black character to save the white ensemble and less likely to sacrifice the white character

to save the black ensemble. In other words, Study R found people weigh white people's lives more than black people's lives in trolley cases. Study R would be good evidence that we should not trust moral intuitions involving race because we generally agree that race is not one of the morally relevant features in trolley cases.

To illustrate how the experimental attack works, imagine another possible world, call it R-World, where analytic philosophy is openly racist.⁹ Philosophers in R-World take it as a plausible and defensible claim that white lives have value that non-white lives do not. In R-World, Study R would therefore be no cause to worry about the epistemic value of intuitions. If anything, Study R would confirm to philosophers in R-World that philosophical intuitions are on the right track because in R-World, philosophers take race to be a truth-relevant feature of the case.¹⁰ Therefore, even if it is ultimately critical of armchair philosophy, evaluative experimental philosophy and the experimental attack is embedded in the practice it is critiquing.

1.4 Intuition Denial

The works being introduced in this chapter deny claim 3 of the experimental attack. By arguing that intuitions do not play a role in philosophy, such INTUITION DENIERS¹¹ can undermine the motivation behind doing evaluative experimental philosophy in the first place. Evaluative experimental philosophy – or at least case-based versions of it – test people's intuitions. Intuition deniers argue that intuitions are not used by philosophers, so evaluative experimental philosophy is epistemically uninformative (Cappelen, 2012, 2014a; Deutsch, 2010, 2015, 2016; Ichikawa, 2016; Ichikawa & Jarvis, 2013; Williamson, 2007).

Existing versions of intuition denial take three very different forms. Ichikawa (2016);

⁹ You can also imagine I am writing this during the Enlightenment.

¹⁰ Thanks to Joseph Bowen for discussion about this argument.

¹¹ This term is from Nado (2016).

Ichikawa and Jarvis (2013) argue – to massively oversimplify – that the focus on intuitions ignores the prior rational pressure people have to believe *a priori* truths. Williamson (2007) and the first part of Cappelen (2012) argue metalinguistically that “intuition” is either meaningless or not an epistemically meaningful category. I address this metalinguistic strategy in Chapter 7. For present purposes, I focus on an exegetical strategy employed first in Deutsch (2009) and Deutsch (2010) then developed and defended in Cappelen (2012) and Deutsch (2015). Deutsch and Cappelen employ the same general strategy and draw the same general conclusions: analyses of original presentations of thought experiments reveals they do not rely on intuitions to reach their conclusion.

Intuition denial has proven extremely influential to debates about intuitions. While intuition denial has rarely been defended outright, a number of observations and arguments from works of intuition deniers have been adopted by other metaphilosophers. As I discuss in Chapter 3, many authors have embraced Deutsch’s innovation of using exegesis, albeit typically in opposition to intuition denial (Chalmers, 2014; Chudnoff, 2017). Similarly, many metaphilosophers have dropped intuition-talk from their work on the epistemology of philosophy (cf. Ludwig, 2007, 2013). Much of this removal of intuition-talk seems to be because many metaphilosophers now acknowledge that intuition-talk is distracting and unnecessary to discuss the epistemic import of experimental philosophy rather than being an acknowledgement of defusal of the experimental attack (see especially Colaço & Machery, 2017; Machery, 2017).

The view put forth by both Deutsch and Cappelen can be described with the admittedly unwieldy label of EXEGESIS-BASED INTUITION DENIAL. Rather than arguing from epistemic points or metalinguistic points, exegesis-based intuition denial argues against the role of intuitions in philosophy by careful examinations of philosophical works. It turns out, they correctly observe, that when you read the presentation of thought experiments in works like Gettier (1963), Lehrer (1990), or Thomson (1976a), there is no

mention of intuitions. More contentiously, Deutsch and Cappelen also claim that in these presentations there is no sign that the authors relied on intuitions to reach their conclusions. Instead, they argue that when we look at the texts all we find are arguments for verdicts about the thought experiments.

The two defenders of exegesis-based intuition denial, Deutsch and Cappelen, differ in a number of ways, but for purposes of exposition, I focus on the arguments in Deutsch (2015).¹² For one, Deutsch focuses on cases central to recent debates on the epistemology of philosophy, namely Gettier cases and Kripke's Gödel case. Cappelen in contrast, discusses a number of notable thought experiments (e.g., Thomson's violinist (1976a) and Perry's essential indexical cases (1979)), but he does not choose thought experiments that are widely discussed in metaphilosophical debates or studied by experimental philosophers. Moreover, Deutsch's exegetical project gets a book-length treatment whereas Cappelen's does not. This means Deutsch's discussion of the Gettier case in particular is better developed than any other exegetical analysis by either of the two authors.

The focus on Deutsch does not change the fact that I am targeting exegesis-based intuition denial in general. The problems I raise in the next three chapters are problems any argument of the form will face. Deutsch (2015) is merely being used as the strongest extant version of the strategy, and I flag places where Cappelen's view differs significantly from it.

1.4.1 Introducing Deutsch

Intuition denial runs against the prevailing self-conception by analytic philosophers that intuitions are key sources of evidence in philosophy. This self-conception is reflected in

¹² Joachim Horvath has also recently defended exegesis-based intuition denial in a talk. I do not discuss it here because I have not seen a manuscript. From what I remember of the talk and from our discussion of a published version of Chapter 2 (Landes, 2020), he *may* have a response to the argument presented there. I do not believe he has a response to the argument in Chapters 3 and 4.

the centrality of intuitions in experimental philosophy and contemporary debates in the epistemology of philosophy. Whereas Cappelen (2012, 105-110) diagnoses this mistake as a combination of a viral spread of intuition-language and influence from rationalists in the 1990s (Bealer, 1996; BonJour, 1998),¹³ Deutsch diagnoses the mistake on insufficient attention to an ambiguity in “intuition”.

Deutsch points out that mental state terms like “judgment”, “belief”, and “intuition” can refer to the mental state itself (the state sense) or the content of the mental state (the content sense). To see the distinction between the state and content sense of mental state terms, consider the sentence “Rebecca judged that she could not jump the gap”. In the state sense, “judged” refers to the process by which Rebecca arrived at her conclusion. The state of judging for Rebecca is a diachronic mental state that calculated distance, ability, surface, etc. to arrive at a conclusion about her abilities. In the contrasting content sense, “judged” in “Rebecca judged that she could not jump the gap” refers to the conclusion drawn, *that the gap is too wide to jump*. While both the state and the content of the state are closely related, what is true of one is not true of the other, and the two should therefore be distinguished.

Deutsch argues that the experimental attack on intuitions has been motivated, in part, by confusion about the content/state ambiguity in “intuition”. In particular, the ambiguity causes a confusion about the relationship between evidence and intuitions. This confusion leads, Deutsch contends, to unreflective endorsement of:

Evidence Claim (EC): Many philosophical arguments depend on treating intuitions about thought experiments and cases as evidence (Deutsch, 2015, 34)

According to Deutsch, EC has a true and false reading depending on whether “intuition” is understood in the state or content sense. The true, content reading of EC grants

¹³ Hintikka (1999) instead blames the influence of Chomsky and early modern rationalists.

that intuitions exist and that their content is often *coextensive* with content taken as evidence. This true reading does not, however, claim the mental *state* is used as evidence:

(ECC): Many philosophical arguments treat the contents of certain intuitions as evidence for or against other contents (e.g., the contents of more general principles) (Deutsch, 2015, 35)

ECC makes no comment on the epistemic status of the content of intuitions. It merely states that certain propositions are used in arguments and that sometimes we intuit those propositions. ECC does *not* claim a proposition's use in an argument is connected to it being intuited. Similarly, I may desire P to be true, have independent reason to think P is true, and use P as evidence without my desire being used as evidence that P.

Notice that by granting ECC, Deutsch accepts that intuitions are a legitimate category of mental states. Whereas Cappelen thinks “intuition” is either meaningless or synonymous with “pre-theoretic judgement” or “snap judgment”, Deutsch is not a deflationist. Nevertheless, Deutsch purposely says very little about what intuitions are. Deutsch instead opts to use a “no-theory” theory of intuitions (see also Malmgren, 2011). The no-theory theory, Deutsch argues, sidesteps a number of contentious issues in the debate over intuitions. Deutsch does not need to commit to whether intuitions are, for example, inferential, perceptual, beliefs, inclinations to believe, or *sui generis* mental states. Nor does he need to commit to whether intuitions have content that is conceptual, psychological, or modal.

The no-theory theory is demonstrative and assumes a theory-neutral understanding of the phenomena. According to Deutsch's no-theory theory, intuitions are *those things*, namely the mental state associated with thought experiments with the content *that Smith does not know* or *that it is permissible to push the man in front of the trolley*. The no-theory theory relies on philosophers' understanding of what “intuition” means based on

paradigmatic thought experiments like trolley cases or Gettier cases.¹⁴ Nothing more specific is needed in order to discuss the evidential role of intuitions, Deutsch argues, because anything more specific and controversial would expose his discussion to purportedly unrelated criticism.

Setting worries about the no-theory theory aside for the purposes of discussion of Deutsch and returning to the Evidence Claim (EC), even though Deutsch thinks intuitions exist and that they lead to beliefs (2015, 75 fn2), he argues they are not used as evidence for positions by philosophers. Philosophers think otherwise, he claims, because of mistaken endorsement of the false, state reading of EC:

(ECS): Many philosophical arguments treat the fact that certain contents are intuitive as evidence for those very contents (Deutsch, 2015, 35)

According to ECS, Gettier used the fact that he and others intuit that “Smith does not know” as evidence for his conclusion that the JTB analysis of knowledge is false. If the claim was not intuitive, ECS falsely holds, then Gettier would have had much less evidence for his conclusion (or perhaps no evidence at all).

Deutsch argues that the conflation of the false ECS and the almost trivially true ECC motivates the MYTH OF THE INTUITIVE (MoI):

As a matter of fact, intuitions about thought experiments are essential evidence for some philosophical positions (2015, xiv-xv)

MoI specifies the weight intuitions play in philosophy. ECS is silent on the point, and ECS is compatible with intuitions playing a small, incidental role in some debates. MoI, in contrast, claims that philosophers hold certain positions because they have certain intuitions as evidence. More specifically (because MoI is understood in the state sense of “intuition”) philosophers hold some positions on the grounds that particular propositions

¹⁴ I raise the worry that this problematically underdetermines the referent of “intuition” in Section 7.6.2.

were intuited. As his title suggests, Deutsch spends most of *Myth of the Intuitive* mustering evidence against MoI with the hope that doing so will undermine the skepticism raised by experimental philosophy.

1.4.2 Defusing the Experimental Attack

Deutsch's main attack on MoI and ECS takes the form of careful methodological exegesis of the original presentations of thought experiments. In particular, he focuses on Kripke's Gödel case (1981) and Gettier's 10 coin case (1963). Deutsch argues that, instead of taking intuitions as evidence for premises, careful analysis of the texts show that Gettier and Kripke are defending their verdicts by argument instead. This is distinct, for example, from Williamson's claim that thought experiments are themselves disguised arguments (2007). On Deutsch's view, original presentations of thought experiments present genuine – as opposed to intuitive – counterexamples to generalizations, and insofar as the genuineness of the counterexamples are defended, they are defended by arguments. This is not to say that Deutsch claims that philosophy is arguments all the way down. Deutsch rightly points out works of philosophy have to stop eventually, and he thinks it is an open question whether the foundational claims of philosophy rest on intuitions. Nonetheless, Deutsch contends that intuitions are not in the places typically thought, and by extension, intuitions are not in the places experimental philosophers have found truth-irrelevant sensitivities.

Denying MoI this way is meant to insulate philosophy from experimental criticisms of intuitions' role as evidence in philosophy. To reiterate, the experimental attack has the following four steps:

1. According to experimental findings, intuitions are affected by P
2. P is truth-irrelevant
3. Philosophy involves reconciling intuitions with principles

4. Therefore, philosophy is in bad epistemic standing

Taking step 1 (the empirical claim) and step 2 (the claim of truth-irrelevance) together, experimental data of truth-irrelevant sensitivities in intuitions motivates the following skepticism:

Intuition Skepticism: Intuitions about thought experiments are not good guides to the truth

Intuition skepticism by itself is not a problem for philosophy for the same reason there is space for Deutsch to hold that intuitions have true contents without being committed to their use as evidence in philosophy. On the traditional metaphilosophical picture of intuitions, intuitions are an epistemic means to discover truths about the targets of philosophy – e.g., knowledge, goodness, and rationality (or the concepts thereof). The experimental attack therefore moves from skepticism about the epistemic status of intuitions (steps 1 and 2) to a more pernicious skepticism about the positions philosophers hold (step 4):

Position Skepticism: Philosophers are not justified in holding (some or all) philosophical positions they believe¹⁵

Intuition skepticism does not entail position skepticism, and the experimental attack needs a bridge between the two. This is step 3 in my version of the experimental attack. Deutsch interprets the bridge to be MoI, and Cappelen interprets the bridge to be Centrality, the slightly weaker claim that “philosophers rely on intuitions as evidence (or as a source of evidence) for philosophical theories” (2012, 2).

Deutsch and Cappelen set out to deny the Myth of the Intuitive and Centrality, respectively, in order to defuse the experimental attack. It does not matter if intuitions are wildly flawed sources of evidence if no one uses them as evidence, so a denial of the Myth of

¹⁵ This view is called “philosophical skepticism” in a different context by van Inwagen (1996).

the Intuitive or Centrality, according to Deutsch and Cappelen, insulates philosophy from experimental studies of reactions to thought experiments. Instead, Deutsch and Cappelen claim classic presentations of thought experiments rely on arguments in order to establish the verdicts of the thought experiments. The main propositions at the heart of thought experiments – whether it is permissible to pull the lever, whether Smith knows, whether “Gödel” refers to Gödel – are argued for by the authors of the thought experiments.¹⁶

Next, in Chapter 2, I show why this does not actually avoid position skepticism.

1.5 Conclusion

In this chapter I introduced the plan for the dissertation and introduced some necessary background information. Over the next six chapters, I explore the relationship between the subject matter of philosophy and the discipline of philosophy. I argue that philosophers interested in the epistemology of philosophy need to draw upon a wide range of empirical evidence.

The bulk of this introductory chapter has been committed to setting up the discussion of exegesis-based intuition denial, which I focus on in the next three chapters. Section 1.2 introduced experimental philosophy, Section 1.3 introduced the four-step experimental attack, and Section 1.4 introduced the intuition denial of Deutsch and Cappelen. With this out of the way, I now turn to two distinct criticisms of exegesis-based intuition denial. Next chapter, I argue that the view is committed to position skepticism. I use this to

¹⁶ Notice that the claim that arguments defend verdicts is different than the claim that thought experiments are used to support arguments. This latter claim is rather trivial. Thought experiments are not free-floating vignettes but are used to establish broader claims. Gettier, for example, uses the verdict “Smith has justified true belief without knowledge” (or something similar) to establish the conclusion “it is not the case that JTB is knowledge” (or something similar). Intuition deniers such as Deutsch and Cappelen are instead focusing on how authors defend the verdict of thought experiments. In the case of Gettier, this is the claim that that Smith lacks knowledge but has justified true belief. Verdicts such as this, Deutsch and Cappelen claim, are supported by arguments in the original text. (For discussion about what exactly the content of the Gettier intuition is, see Malmgren (2011); Williamson (2007).)

argue that historical facts about how works of philosophy affect debates are important for understanding the epistemology of philosophy. After that, in Chapter 3 and 4, I argue that there is no motivation for the method of textual exegesis in the place. This allows me to defend the importance of studying empirical facts about how people *read* – not write – works of philosophy. From there, I leave criticism of Deutsch and Cappelen aside to explore the importance of lexicography to metaphilosophical projects, including the study of intuitions.

Chapter 2

The Threat of the Intuition-Shaped Hole

In Chapter 1, I introduced the dialectic for this chapter. Experimental philosophers purport to raise worries about the epistemic standing of philosophy through experimental data on intuitions. This involves a four-part argument that ties together experimental data, philosophical judgements about truth-irrelevance, and methodological facts about the use of intuitions by philosophers. Intuition deniers reject this argument by denying that philosophers use intuitions in an epistemically meaningful way. In particular, exegesis-based intuition deniers reject the methodological step of the experimental attack – and the resulting skepticism about philosophy – by carefully examining the methods described in original presentations of thought experiments. In this chapter, I argue that this strategy of avoiding skepticism is ultimately unsuccessful.

Previous replies to intuition denial have focused on the descriptive inadequacy of intuition denial. Climenhaga (2017) argues abductively from facts about philosophers' behavior to the conclusion that intuitions do play justificatory role, and Brown (2017) argues that Deutsch's view cannot explain why epistemologists working in the 1960s and 1970s understood the Gettier verdict as justified. More commonly, critics argue that

that misunderstandings about how intuitions are thought to work have led Deutsch and Cappelen to fail to establish the conclusion that intuitions play no epistemic role in philosophy (e.g., Baz, 2016; Chalmers, 2014; Chudnoff, in press; Weinberg & Alexander, 2014).¹

This chapter explores a different and unexplored issue inherent in exegesis-based intuition denial. Even if authors like Deutsch and Cappelen provide a descriptively adequate and intuition-free account of philosophy, they have not saved philosophy from skeptical worries.² Given that the sorts of thought experiments under consideration have been generally persuasive, exegesis-based intuition deniers only prevent skepticism about philosophy if they can explain how beliefs caused by considering thought experiments are justified. Deutsch or Cappelen do not have the resources to do so. Instead, while exegesis-based intuition denial may prevent the claim that philosophers use problematic intuitions as evidence, it is still committed to the claim that philosophers widely adopt philosophical positions on bad grounds.

In Section 2.1, I argue that to properly block skepticism about philosophical practice, intuition deniers have to provide an adequate epistemic replacement for intuitions. Then in Section 2.2, I put forth what I take to be Gettier's argument against the JTB account of knowledge – if there is one. Section 2.3 denies that *any* plausible intuition-free reading of Gettier's argument will justify beliefs in the falsity of the JTB analysis of knowledge. Section 2.4 then argues against the textual evidence that Gettier is in fact arguing for his conclusion as opposed to intuiting it. I end the chapter in Section 2.5 by generalizing this argument to thought experiments beyond Gettier cases and arguing the points in this chapter represent a general problem for exegesis-based intuition denial.

¹ See Deutsch (2016) and Cappelen (2014a, 2014b) for responses to past criticisms.

² There are hints of this approach in (Nado, 2016, §4) and Colaço and Machery (2017).

2.1 The Threat of the Intuition-Shaped Hole

As discussed in Section 1.4.2 the intuition denial of Deutsch (2015) and Cappelen (2012) attempts to block the experimental attack by blocking the inference from skepticism about intuitions to skepticism about philosophical beliefs. To see why removing intuitions from notable works containing thought experiments will not save philosophy from position skepticism, consider a concession made by both Deutsch and Cappelen. While developing an intuition-free version of Lehrer's (1990) Mr. Truetemp thought experiment, Deutsch considers the objection that without intuitions, Lehrer's argument is not particularly strong. To this possibility Deutsch concedes,

The question is not whether Lehrer's arguments for his judgment about the case are good arguments. The question is rather whether intuitions about the Truetemp Case play the evidential role that Swain et al. (2008) assign to them. (2015, 113)

Cappelen makes a similar point while considering his intuition-free exegesis of Thomson's ailing violinist (1976a). In a footnote, Cappelen says that his exegesis's failure to establish Thomson's stated conclusion "simply shows that Thomson's argumentative strategy is problematic" (2012, 156 fn13).

Both passages overlook an unnoticed and possibly fatal problem with exegesis-based intuition denial. The works studied by metaphilosophers, such as Gettier (1963), Thomson (1976a), and Lehrer (1990), are widely considered by philosophers to be successful attacks on their respective targets. If exegesis-based intuition deniers cannot explain why these authors' verdicts are good conclusions to draw without intuitions, then exegesis-based intuition deniers have not escaped position skepticism. They are forced to concede that philosophers have historically accepted philosophical positions based on bad arguments. This upshot is the same as the experimental attack – philosophers have reason to think that they are not justified in believing the philosophical positions they actually believe.

Therefore, in order to insulate philosophy from position skepticism, exegesis-based intuition denial needs to replace intuitions with a different, epistemologically satisfactory story. This story must explain why countless philosophers have, on the basis of passages containing thought experiments, justifiably accepted that the descriptive theory of reference is unsatisfactory and that justified true belief is not sufficient for knowledge. Otherwise, exegesis-based intuition denial fails its central motivation.

Examining every exegetical account developed by Deutsch and Cappelen would be a monumental undertaking, so the rest of this chapter will focus specifically on Deutsch's analysis of Gettier's 10 coins case. Deutsch is more explicit than Cappelen about what exactly he thinks the arguments behind thought experiments are, developing a sophisticated and compelling intuition-free account of Gettier's rejection of the justified true belief account of knowledge. Moreover, rather than focusing on thought experiments peripheral to metaphilosophical debate, as Cappelen does, Deutsch focuses on the 10 coins case – a thought experiment central to experimental philosophy and considered an exemplar of the case method.

This chapter explores whether Deutsch has an epistemic replacement for intuitions sufficient to avoid position skepticism. Drawing from both Gettier's original presentation and comments by Deutsch, I develop my own argument-based account of Gettier. From there, I argue that without intuitions, the candidate arguments have gaps that leave Gettier's conclusion unjustified. This shows that, at least without a more robust replacement than he provides, Deutsch does not satisfactorily prevent position skepticism. From there I argue that issues raised below are not idiosyncratic to Gettier's 10 coins case. Instead, the issues raised are individual instances of a global problem facing exegesis-based intuition denial.

2.2 The 10 Coins Case

While Gettier presents two thought experiments, Deutsch focuses on the first of the two, the 10 coins case. Gettier describes a scenario in which Smith hears from his boss that Jones is about to get a big promotion. Smith has also learned that Jones has 10 coins in his pocket. From these two beliefs, Smith forms a third belief, E, that the person who will get the job has 10 coins in their pocket. Because Smith has this information on good authority, Smith is justified in holding E. As it turns out, however, Jones is not getting the job – Smith is. Moreover, Smith has 10 coins in his pocket. Therefore E is true. Nonetheless, Gettier points out, Smith does not know E, even though he believes E, is justified in believing E, and E is true. From this and a second similar case, the Brown in Barcelona case, Gettier concludes that justified true belief is not sufficient for knowledge.

In order to evaluate whether Gettier is in fact defending his verdict by argument, it is worth trying to explicate what that argument would be. Deutsch gives few clues about what he takes the full argument to be, except that it likely starts as early as Gettier's description of the case (2015, 77 fn4). Starting there, this is the 10 coins case converted almost verbatim from Gettier's original presentation into numbered premises, with careful attention paid to giving as robust an argument as the text allows:

1. **Assumption:** Justified true belief is knowledge.
2. **From Description:** E [the man who will get the job has 10 coins in his pocket] is true.
3. **From Description:** Smith believes that E is true.
4. **From Description:** Smith is justified in believing that E is true.
5. **From 2, 3, and 4:** Smith has a justified true belief that E.
6. **From Description:** E is true in virtue of the number of coins in Smith's pocket.
7. **From Description:** Smith does not know how many coins are in Smith's pocket.
8. **From Description:** Smith bases his belief in E on a count of the coins in Jones' pocket, whom Smith falsely believes to be the man who will get the job.
9. **From ?:** Smith does not know E.
10. **From 4 and 9:** Smith has a justified true belief that E and does not know E.
11. **From 1 and 10:** \perp
12. **From 1 and 11:** Justified true belief is not sufficient for knowledge.

The question at the heart of this chapter is whether this argument, or something sufficiently close, is justified without intuitions. There are two distinct ways to understand this question. First, we can evaluate how Gettier justifies the verdict of the 10 coins case, a key moment in his larger argument. In particular, does Gettier satisfactorily defend the verdict that Smith does not know? The locus of this first question is premise 9 of the above formalization – Gettier's verdict about the thought experiment. Intuition-based readings of Gettier hold that premise 9 is justified via an intuition while Deutsch contends it is defended by an argument.³

In contrast, we can evaluate whether Gettier provides a good argument for step 12, his

³ I assume this argument is deductive. Gettier does not seem to intend to defend the verdict abductively or inductively. Abductive inferences try to explain phenomena, and Smith's lack of knowledge is not, for example, invoked to explain why Smith believes that the man who will get the job has 10 coins in his pocket. If anything, as discussed in Section 2.4, the verdict is the thing that needs to be explained. Similarly, the argument to 9 is not obviously inductive because there are not any generalizations involved. This is a single verdict about a single instance that does not depend on extrapolating from past experience (see, however Colaço & Machery, 2017, §5).

conclusion. This requires that we evaluate Gettier's larger paper-level argument against the JTB analysis of knowledge. Specifically, how does Gettier satisfactorily defend the claim that justified true belief is not sufficient for knowledge? The answer to this question will ultimately involve a *reductio ad absurdum*. At the start of the paper, Gettier sets out the assumption that justified true belief is sufficient for knowledge. As the paper continues, Gettier provides two counterexamples to the principle: Smith's lack of knowledge in the 10 coins case and the Brown in Barcelona case. Gettier ends the paper by stating that he has shown that justified true belief is not sufficient for knowledge. By laying out an assumption and (implicitly) deriving a contradiction, Gettier deductively rejects the JTB analysis.

Both notions of whether Gettier makes a good argument are crucial for the purposes of this chapter. In order to provide a sufficiently good argument without intuitions, intuition deniers need a satisfactorily account of how Gettier defends Smith's lack of knowledge (step 9) *and* how Gettier rejects the JTB analysis (step 12). A failure to explain the former means that Gettier's verdict that Smith does not know is not justified. A failure to explain the latter means that Gettier's rejection of the JTB analysis is not justified. While an account of how the verdict is defended is necessary to provide an account of how Gettier rejects the JTB analysis, it is not sufficient. Without both, exegesis-based intuition deniers are forced to admit countless philosophers have rejected the JTB analysis because of a bad argument, thereby leading to the very skepticism they hope to avoid.

2.3 Defending the Verdict Without Intuitions

Most of the above formalization of Gettier follows a straightforward form. Premises 2, 3, and 4 are stipulated by Gettier during the case and entail premise 5. Similarly, premises 6, 7, and 8 describe stipulated features of the case. The sub-conclusion, premise 10, is entailed by 5 and 9, which in turn entails the contradiction and the negation introduction

of the assumption – premise 1 – in the argument’s final step.

This leaves the verdict, 9, that Smith does not know E. Either 9 is assumed as a premise (and justified externally to the argument), or it follows from other premises. Deutsch seems open to both possibilities, and both possibilities are reflected in Gettier’s language. Consider the passage in which Gettier first gives his verdict: “But it is equally clear that Smith does not know that [E] is true; for [6, 7, and 8]” (1963, 122). There are signs of Gettier taking *both* options in this passage. “Equally clear” suggests Gettier is appealing to obvious external justification, while “for” suggests Gettier takes 9 as following from 6, 7, and 8.

While Deutsch considers both options, each poses unique problems for the strategy of intuition denial. This raises a fork for Deutsch and exegesis-based intuition denial more generally. Either the verdict is defended with external justification – and it is not clear what that could be besides intuitions – or the verdict is defended through other premises in the argument – in which case the argument is not valid. I look at each possibility in turn.

2.3.1 Externally Justified Verdicts

In this subsection, I examine whether 9 can be taken as a premise justified externally to the argument. Unimpressed by Gettier’s intuition-talk of “equally clear”, Deutsch suggests that Gettier does not see the need to defend 9 because 9 is too obvious to warrant a defense:

So when Gettier claims that it is “clear” that his Smith character does not know that the man who will get the job has ten coins in his pocket, he is not adducing evidence for the claim, let alone evidence that consists in his or other people’s intuitions. He is instead pointing to the ease with which one can recognize that claim is true. (2015, 81)

In other words, Gettier does not appeal to intuitions because he does not have to. The case is so clear and easy to understand that Gettier does not have to defend Smith's lack of knowledge. Intuition deniers cannot just excise intuitions from thought experiments by appealing to external justification, however. Doing so raises a serious issue about the judgement's epistemic basis. It is not clear how, if not from intuitions, we can come to know about the targets of philosophy when using thought experiments.

This is not to say Deutsch does not have a purported story to explain how we know the Gettier verdict without intuitions. Deutsch, like Cappelen, is an extra-mentalistic about philosophy; he takes the target of philosophical inquiry to be real, mind-independent entities – in contrast to psychological concepts or linguistic entities (Deutsch, 2015, 48-50). Thus, Smith's lack of knowledge is a fact about the world, not a fact about what we think or how we speak. Building upon his extra-mentalism, Deutsch argues that our knowledge of counterexamples is unproblematic and non-psychological. He asks us to imagine that his daughter claims that she put away all of her toys. In order to provide a counterexample to his daughter's claim, he merely needs to point to things in the world, namely, all of the toys that are still on the floor. There is nothing mysterious about having a toy as a counterexample, and there is certainly nothing in it that relies on intuitions. Deutsch knows his daughter is wrong because there is an object he perceives, a toy on the floor. A similarly straightforward story can be told for some counterexamples in science.⁴ To pick an artificially simple example, how did ornithologists come to know that not all swans are white? They discovered black swans.

Matters are significantly less simple in philosophy. It is not clear that we can tell the same story about seeing knowledge (or goodness, justice, justification, etc.) that we can tell about seeing a toy or a black swan. Even if you take the line that judgements about knowledge are versions of the same mundane categorization processes that allow us to

⁴ Certainly not all counterexamples in science have a straightforward epistemology. Galileo was able to overthrow the impetus theory in physics with a thought experiment. See Gendler (2000) for an overview.

pick out toys and swans (e.g., Machery, 2017; Williamson, 2007), telling a story about how we calibrate our categorizations about targets of philosophy will be significantly more difficult than everyday objects (Baz, 2016; Cummins, 1998). Intuitions are exactly the sorts of mental states that are thought to give us access to the targets of philosophy. Intuitions are often evoked to explain how we can grasp that certain philosophically relevant concepts – such as knowledge – apply to a situation (e.g., Kauppinen, 2007; Ludwig, 2007), and intuitions are sometimes described as a perception-like process that gives us access to philosophical knowledge similarly to how eyesight gives us access to the visual world (e.g., Bengson, 2015; Chudnoff, 2013b). Deutsch does not explain how “the ease by which we recognize” that Smith does not know is not a tacit appeal to intuitions of some variety, such as perceptual intuitions. Perhaps there is way to explain this while maintaining exegesis-based intuition denial, but it cannot be done by merely appealing to mundane counterexamples like toys or swans.⁵

2.3.2 Invoking Principles

As mentioned above, intuition deniers have two ways to justify verdicts about thought experiments – either the verdict is justified externally, or it is entailed by other parts of the argument. As just discussed, the first fork leads to an issue about what provides justification for the target belief, an issue that Deutsch does not have a satisfactory response to. I now consider the second fork. Intuitions are not required to justify beliefs if the verdicts under consideration (in this case premise 9) are entailed by other premises in the argument and those other premises do not need intuitive justification. This is a position Deutsch explicitly takes, arguing that Gettier’s use of “for” after giving the verdict and before describing the features in premises 6, 7, and 8 indicates Gettier takes the verdict to follow from these premises (2016, 75).

⁵ Weinberg and Alexander (2014) raises similar points.

This is not enough to show how the verdict is justified, however, because the sub-argument from premises 6, 7, and 8 to 9 does not have a deductively valid form. A principle is required to move the argument from the specific features of the case discussed in premises 6, 7, and 8 to a conclusion about knowledge (premise 9). This would be something of the form “if x , y , and z , then P does (not) know”. The above argument is missing such a principle, but perhaps adding one improves Gettier’s argument and, by extension, helps exegesis-based intuition deniers. In discussion of Gettier’s other thought experiment, the Brown in Barcelona case, Gettier says that Smith’s belief is true in that instance by the “sheerest coincidence” (1963, 123). Since the cases are related, Gettier likely saw them as sharing essential features, and therefore discussion about one thought experiment is relevant to the other. Deutsch’s candidate principle to exclude “sheerest coincidence” is:

Epistemic Luck (EL) If S ’s belief that p is (also) knowledge that p , then S ’s belief that p is not luckily true. (Deutsch, 2015, 88)

Epistemologists had argued before Gettier’s paper that luck cannot give rise to knowledge, so perhaps the principle was well-known enough that Gettier could have merely gestured to it (Deutsch, 2015, 89 fn14). Accordingly, here is a shortened version of the purported Gettier argument, adding a transposed version of EL and the premise that Smith is lucky:

1. **From Description:** Smith has a justified true belief that E.
2. **From Description:** E is true in virtue of the number of coins in Smith's pocket.
3. **From Description:** Smith does not know how many coins are in Smith's pocket.
4. **From Description:** Smith bases his belief in E on a count of the coins in Jones' pocket, whom Smith falsely believes to be the man who will get the job.
5. **From ?:** Smith's belief in E is luckily true.
6. **Assumption (EL):** If S's belief that p is luckily true, then S's belief that p is not knowledge that p.
7. **From 5 and 6:** Smith does not know E.
8. **From 1 and 7** Smith has justified true belief that E and does not know E.

This argument, if a proper analysis of Gettier, removes intuitions about knowledge from Gettier's larger reductio. It is not clear, however, that it removes the need for intuitions entirely. The added principle merely moves the problem facing Deutsch from the application of knowledge to the application of luck. The same challenge facing the verdict about knowledge can be raised here for the verdict about luck. While there is a principle stipulating that luck is inconsistent with knowledge, there is no principle that formally connects the features of the case to the conclusion that Smith's judgment is lucky. Therefore, if we want to understand the verdict "Smith is lucky" as following from earlier in the argument, adding EL requires adding yet another principle to determine when someone is lucky. At various points, Deutsch suggests he would be unmoved by this issue. Papers are finitely long, and arguments have to stop somewhere (Deutsch, 2015, 123-124). Therefore, Gettier's argument might as well bottom out at luck. Moreover, Deutsch might be willing to take the other side of the fork and take "Smith is lucky" to be intuited. Unlike Cappelen, Deutsch does not categorically deny that intuitions have a role in philosophy. Rather, Deutsch claims that intuitions are not essential evidence and are not in the places experimental philosophers and methodologists have assumed they are (2015, 122-127).

Alternatively, Deutsch can avoid the fork entirely by claiming the phrase “sheerest coincidence” is working double duty, both invoking a principle of luck and stipulating that Smith is lucky.⁶ In this case, intuitions are not needed to justify the verdict that Smith is lucky, and the argument is not lacking any principle to connect features of the case with the rejection of the JTB account. Nonetheless, any reading of the Gettier argument that includes EL prevents exegesis-based intuition deniers from being able to give a satisfactory account of Gettier’s rejection of the JTB analysis.

2.3.3 A New Problem

Bringing in EL is meant to explain how Gettier justifies the conclusion that Smith does not know without appealing to intuitions about knowledge, but doing so undermines Gettier’s larger reductio. If Gettier defends Smith’s lack of knowledge via a principle of luck, then he cannot justify the conclusion that JTB is not sufficient for knowledge. Because such an argument requires assuming both the truth of EL and JTB, we no longer truly have a reductio against JTB by itself. Rather, it creates a reductio against the *conjunction* of JTB and EL:

⁶ Thanks to Justin Snedegar for this suggestion.

1. **Assumption (JTB):** Justified true belief is knowledge.
2. **Assumption (EL):** If S's belief that p is luckily true, then S's belief that p is not knowledge that p.
3. **From Description:** Smith has justified true belief that E.
4. **From Description:** Smith's belief that E is luckily true.
5. **From 1 and 3:** Smith knows E.
6. **From 2 and 4:** Smith does not know E.
7. \perp

This argument does not reveal whether EL or JTB is false. It merely demonstrates that they cannot both be true. If this were Gettier's intention, the thought experiment would be what Cappelen calls a "fact focuser" (2012, 133-134) – a thought experiment that draws our attention to something we had previously not considered. If the Gettier case is merely a fact focuser, then the case draws our attention to an unnoticed tension between JTB and EL without being a counterexample to either. Accepting that many thought experiments are merely fact focusers as opposed to counterexamples, a move Cappelen embraces, would give intuition deniers an easy way to eliminate intuitions from accounts of how thought experiments work, since, as the above argument demonstrates, it does not require the sort of justification intuitions are thought to provide. Such a move would blatantly move the goalposts, however. Gettier explicitly concludes that his cases show that JTB is not sufficient for knowledge. More importantly, *other epistemologists* have taken Gettier as showing JTB is not sufficient for knowledge, and to avoid position skepticism, any intuition free exegesis must account for this.

Given that exegesis-based intuition deniers need to account for how Gettier rejects the JTB analysis, if both EL and JTB are invoked in Gettier's argument, then Gettier (and his readers) need some justification for weighing EL more than JTB when the two conflict. As Gettier discusses at the beginning of his paper, the JTB analysis of knowledge

is a venerable theory, with versions dating back to Plato. If anything, the JTB analysis deserves the benefit of doubt over a principle of epistemic luck.⁷

In order to arrive at the desired verdict that Smith does not know, we need some principle. According to Deutsch, this is the principle of epistemic luck. Including such a principle adjusts the structure of the resulting argument, so we then need some reason to think that we should maintain the principle of luck more strongly than the JTB principle. Nowhere in Gettier's paper is such an argument supplied. Adding EL to Gettier's argument already strains the limits of charitable interpretation given the length of Gettier's original paper, and adding any additional arguments for EL (or any additional premises) would approach wishful thinking on the behalf of exegesis-based intuition deniers. Gettier does not manage to justify rejection of the JTB analysis without appealing to intuitions.

By removing intuitions from Gettier's defense of the verdict, Deutsch cannot explain how Gettier is justified rejecting the JTB analysis. He does not have a way to show that Gettier's argument is a good argument, and thus does not protect epistemologists from position skepticism.

2.4 Abduction in Gettier

In this section, I explore the possible argumentative functions of the phrase "sheerest coincidence". So far, I have followed Deutsch and assumed that "sheerest coincidence" is used to support the verdict that Smith does not know. This runs against the role "sheerest coincidence" is typically thought to play on intuition-based readings Gettier's paper. If Gettier justifies his verdict with intuitions, then "sheerest coincidence" is offered as an explanation of why Smith does not know (Spicer, 2008), rather than support for that verdict.

The abductive reading of "sheerest coincidence" is roughly the following:

⁷ Notice that this issue is independent of how Gettier justifies the premise that Smith is lucky.

1. EXPLANANDUM: Smith has a justified true belief that is not knowledge.
2. Smith's belief is true in virtue of the number of coins in Smith's pocket.
3. Smith does not know how many coins are in Smith's pocket.
4. Smith bases his belief on the number of the coins in Jones' pocket, whom Smith falsely believes to be the person who will get the job.
5. EXPLANANS: The fact that Smith has a justified true belief that is not knowledge is best explained by the fact that the belief is true by sheerest coincidence.

In this presentation, Gettier includes the phrase because he is offering a best explanation for the pre-established fact that Smith does not know. Gettier is not then using the phrase to argue for the claim that Smith does not know.⁸

Deutsch calls the objection that he is misunderstanding the role of “sheerest coincidence” the ORDER OF EXPLANATION OBJECTION (2015, 95-99) since it concerns the order of derivation between the case's verdict and the statement of Smith's luck. He offers up two flawed responses. The first response argues that when we consider what it is like to create the thought experiment instead of consuming it, the abductive reading of “sheerest coincidence” does not make sense. I discuss the distinction between producers and consumers of philosophy and its relationship to exegesis extensively in Chapters 3 and 4, but an initial point can be made here. In his response, Deutsch claims that if phrases like this were abductive inferences about intuitions, we would see intuition-talk in original presentations of thought experiments. This response highlights a tension in exegesis-driven intuition denial between the claim that philosophers are wrong about their methods and the claim that authors accurately represent their methods on paper. I discuss this in Section 2.4.1.

The other response to the order of explanation objection accepts that “sheerest coincidence” is serving an abductive role but denies that this is incompatible with the phrase

⁸ There are other possibilities. The phrase may be playing no argumentative role at all. Instead, “sheerest coincidence” may be an attempt by Gettier to help elicit the intuition that Smith does not know (see Chudnoff, 2017). On this reading, the phrase exists to draw readers' attention to features of the case that will aid in the formation of the correct intuition. I will set this possibility aside.

supporting the verdict that Smith does not know. In other words, Gettier's paper is making both an abductive and deductive argument. In Section 2.4.2, I discuss why this cannot be right. If the phrase is working double duty in an abductive and deductive argument and the deductive argument does not rely on intuitions, then arguments like Gettier's could illicitly bootstrap any plausible philosophical view.

2.4.1 The Issue with Original Presentations

Deutsch's first response to the order of explanation objection hinges on the lack of intuition-language surrounding original presentations of thought experiments. According to Deutsch's reply, Gettier keeps it open at the start of the paper whether Smith knows. Gettier has to, otherwise he would be begging the question against JTB. Instead Gettier needs some reason to defend the conclusion that Smith does not know. Deutsch argues that that reason cannot be intuitions because intuition-talk does not appear in Gettier (1963) or other original presentations of other thought experiments. Instead Deutsch claims that textual evidence reveals Gettier is arguing for the verdict with "sheerest coincidence" instead.

This argument turns on how seriously we should take the language of original texts. This point is worth discussing because it highlights a set of issues inherent to exegesis-based intuition denial. To be able to conclude from exegesis that papers like Gettier (1963) do not rely on intuitions, Deutsch and Cappelen need to be committed to both the following claims:

- We can examine the original texts in order to determine the authors' true methods.
- It is possible philosophers largely do not understand the methods of philosophy.

The second claim, that philosophers can be widely mistaken about the methods of philosophy, enables Deutsch and Cappelen to reject philosophers' intuition-talk. If Deutsch

and Cappelen are right, then decades of discussion of intuition's role in philosophy have been in error. It has been a pretty obvious error, at that. Without assuming the possibility of widespread misunderstandings about the methods of philosophy, Deutsch and Cappelen would have to take all that intuition-talk at face value as evidence for the use of intuitions in philosophy. They do not, instead taking exegetical evidence to outweigh such talk.

The first claim is also a key assumption of exegesis-based intuition denial. If exegesis did not reveal authors' methods, reading texts to analyze their methods would be, at most, a project in self-reflection into one's interpretation of authors' methods. Instead, in using exegetical data to argue that intuitions in fact do not play a role in philosophy, both Deutsch and Cappelen assume that authors' words are accurate representations of authors' methods.

It is not obvious this first assumption is true, however. Because there is often tension between style and accuracy, original presentations of thought experiments might not, at least by themselves, be good sources of evidence for the actual methods of philosophers. Moreover, this is an untested *empirical* claim. Perhaps philosophers' meta-discourse and logical language reflect their own thought processes and the actual workings of their papers, but perhaps not. We need to test the claim to be sure. Speaking anecdotally, I use a handful of thought experiments in this dissertation as argumentative support for various claims. For the most part the thought experiments significantly pre-date or post-date my own arrival to the conclusion the thought experiments are used to defend. Similarly, for reasons discussed in Chapter 7, I have chosen not to use intuition-talk in relation to my own thought experiments. Nonetheless, I think intuitions play an important rhetorical and justificatory role with these thought experiments.⁹

Setting this aside, both claims combined raise the question of whether exegesis-based

⁹ More accurately: the mental states often referred to with "intuition" play an important role. See Chapter 7.

intuition denial is even possible in the first place. Once we accept philosophers can be mistaken about the methods of philosophy, it is no longer clear that we should take the language of authors as a guide to *their* own methods. If Deutsch and Cappelen think philosophers can competently engage in first-order philosophy without understanding their own methodology, why should they also accept that Gettier understands the methodological moves in his paper? For instance, some methodologists have recently argued that abduction is central to philosophy (e.g., Andow, 2016b; Henderson & Horgan, 2013; Williamson, 2016), despite the fact that philosophers rarely use abductive language such as “the best explanation for x is y” or “x accounts for y” (Deutsch, 2015, 96-97). If these authors are right, philosophers have been mischaracterizing their argumentative structures (just as, Deutsch and Cappelen contend, philosophers have sometimes mischaracterized the role of intuitions).

2.4.2 Worries of Circularity

As just discussed, Deutsch’s first response to the order of explanation objection appeals to the apparent use of arguments instead of intuitions in authors’ writing, but the response relies on questionable claims about our ability to infer authors’ methods from texts. Deutsch’s other response to the order of explanation objection contends that it does not matter if “sheerest coincidence” is arrived at abductively since it can play both roles. Epistemic luck can support the conclusion that Smith does not know which can in turn provide abductive evidence that epistemic luck precludes knowledge.

Deutsch’s idea seems to be the following: Imagine we have two premises – that *P is an a* and *all a’s are b’s* – and a deductively valid conclusion – *P is a b*. The same premises can be “reversed” into an abductive argument about the general principle – *all a’s are b’s* explains why P is both an *a* and a *b*. Given the paucity of information I stipulated, it is a stretch to say that the general principle is the best explanation for the fact that some

object has two properties. But imagine we had a large set of objects that are both *a*'s and *b*'s and no *a*'s that fail to be *b*'s. Assuming we have empirical reasons to believe that our set is representative (Henderson & Horgan, 2013, 40-41), we can also abductively infer that the principle that all *a*'s are *b*'s best explains our data. The actual abductive power of the deductively derived conclusion (that *P is a b*) is weak, but this is to be expected – there are countless possible explanations for *P*'s being a *b*.

While this sort of thinking may be behind Deutsch's response to the order of explanation objection, it fails to recognize that we need justification external to the abductive and deductive arguments in order to run the two in tandem. Otherwise, we can recursively bootstrap false beliefs (see also Colaço & Machery, 2017, 414). Imagine we are considering a thought experiment similar to a trolley case in which we are forced to either kill a woman or a man who is otherwise identical. Take the following unsound deductive argument:

1. We have to kill either the woman or the man who are otherwise identical.
2. **SEXISM**: *Ceteris paribus*, a man deserves more moral considerations than a woman.
3. If **SEXISM** is true and we have to kill either a woman or a man who are otherwise identical, then we ought to kill the woman.
4. Therefore, we ought to kill the woman.

Allowing deductive arguments and abductive conclusions to feed into each other allows us to (weakly) abductively arrive at **SEXISM** by taking the above conclusion as an explanandum:

1. **EXPLANANDUM**: If we have to kill either the woman or the man, we ought to kill the woman.
2. As stipulated by the case, the man and the woman are identical in all relevant ways except for being a man and a woman.
3. The best explanation for **EXPLANANDUM** is that, *ceteris paribus*, a man deserves more moral considerations than a woman.

The deductive argument is valid, and the abductive argument gives us evidence that the deductive argument is sound. Examples such as this can be generated endlessly.

To illustrate what goes wrong in the above arguments, consider a slightly different example:

1. Utilitarianism is true.
2. If utilitarianism is true, then I should not kick puppies for fun.
3. Therefore, I should not kick puppies for fun.

Then:

1. EXPLANANDUM: I should not kick puppies for fun.
2. If utilitarianism is true, then I should not kick puppies for fun.
3. The best explanation for EXPLANANDUM is that utilitarianism is true.¹⁰

The example involving puppies should sit better than the example involving sexism even though the two have a parallel form. This is because the apparent correctness (or plausibility) of the puppy abductive argument relies on smuggling in outside evidence. In the case of kicking puppies, we have independent reason to think the principle and verdict under consideration are both correct. Kicking puppies is wrong, and, at least for many people, utilitarianism is a live theoretical option. Therefore, the fact that utilitarianism entails I should not kick puppies is a fact in favor of utilitarianism. If, like in the sexist case, we do not have independent evidence to believe the premises of the two arguments, then the conclusions from abductive and deductive arguments running in tandem do not tell us anything meaningful. As a principle, SEXISM is not worth taking seriously and we do not think we ought to prefer men over women in trolley cases, so the abductive argument, while explaining the hypothetical data, does not give us reason to support SEXISM.

¹⁰ If these arguments sit poorly, consider the same two arguments with a moral principle you find plausible.

Returning to Gettier, if Deutsch is right and the Gettier verdict plays a role in both a deductive and abductive argument, then the situation is closer to the arguments about sexism than the arguments about kicking puppies. Without intuitions, we do not have independent reason to hold the verdict that Smith does not know. We might be able to derive Smith's lack of knowledge from the principle of epistemic luck, but this does not itself provide abductive support for the principle of epistemic luck because we do not have other evidence for the conclusion that Smith lacks knowledge. If anything, because JTB entails it, Gettier's contemporaries had independent reasons to believe that Smith knows. Thus, if Gettier arrives at Smith's lack of knowledge via a principle of epistemic luck, he does not then have independent evidence supporting the verdict to use in an abductive argument. So, unlike what Deutsch suggests, the principle of luck cannot be involved in both an abductive and deductive argument, and his second argument against the order of explanation objection fails.

2.5 Other Thought Experiments

Throughout this chapter, I have argued that removing intuitions from thought experiments via exegesis creates as many problems as it solves. By eliminating experimentally-driven skepticism about the beliefs philosophers hold, exegesis-based intuition denial introduces a new form of skepticism. So far, discussion in this chapter has focused on Gettier's 10 coins case. In this section, I consider how my arguments related to Gettier (1963) generalize to other thought experiments.

Limiting the above arguments to Gettier cases is itself a worrisome challenge for exegesis-based intuition deniers. While there is some controversy about how widespread the JTB analysis of knowledge was before Gettier (Dutant, 2015; M. Kaplan, 1985), it is important to note the effect Gettier cases have had on subsequent analytic epistemology. Despite an abundance of competing positive accounts of knowledge defended by philoso-

phers of different of methodological stripes, vanishingly few defend the claim that JTB is sufficient for knowledge. The JTB account's lack of popularity after 1963 is striking because the theory has notable virtues. It is commonsensical enough to have survived in one form or another as a serious candidate from Plato to the 20th century. Moreover, for the JTB account's simplicity (especially compared to some contemporary accounts of knowledge), it offers strong theoretical power, able to predict the vast majority of cases of knowledge (Weatherson, 2003). In other words, Gettier's paper has had a huge effect on the positions of analytic epistemologists. The above analysis demonstrates that Deutsch and Cappelen do not have an avenue for justifying Gettier's conclusion that JTB is not sufficient for knowledge, and thus they are committed to the conclusion that the contemporary search for theories of knowledge have been largely influenced by an unjustified belief. If Deutsch and Cappelen want to protect analytic epistemology from the experimental attack, this is not the way to go about doing it.

This leaves the question of how many philosophical beliefs exegesis-based intuition deniers are committed to skepticism about. I have more interesting things to do in this dissertation than re-run the above Gettier exegesis on other thought experiments. Nevertheless, it will be an uphill battle for exegesis-based intuition deniers to escape the skeptical challenge raised here with other thought experiments. They will need to either show that the arguments are good arguments without intuitions or show that the verdicts of the thought experiments are not used as justification for arguments in the work. Neither will be easy for the majority of influential analytic thought experiments.

In this section, I explain why exegesis-based intuition denial will in general struggle to provide justification for works' arguments. Discussion here is split between three different types of argumentative roles that capture a large majority of "classic" analytic thought experiments. The distinction is between thought experiments used to establish counterexamples, thought experiments used to establish positive support for a theory, and thought

experiments used as illustrations for principles.¹¹ These categories are not mutually exclusive. In a single paper a thought experiment can be used as a counterexample, evidence for a positive claim, and an illustration of a theory. Thought experiment play a slightly different justificatory role in each of these uses, and I focus on the difficulties intuition deniers face with the first two roles in particular.

2.5.1 Thought Experiments as Counterexamples

As with Gettier cases, oftentimes thought experiments are used to present counterexamples for various theories or arguments. Classic examples of these include Lehrer's Mr. Truetemp (1990), Thomson's ailing violinist (1976a), and Kripke's Gödel case (1981). In each of these examples, the author has a clear target in mind: a certain form of externalism about knowledge, an anti-abortion argument depending on the priority of a person's right to life, and a consequence of a descriptive theory of reference, respectively. Each author introduces their target, introduces a case, and explains why the case defeats the target. Both Deutsch and Cappelen focus on this final step, arguing that this explanation is the locus of the argument that establishes the verdict for the thought experiment. However, as discussed with the Gettier case, it is not clear whether these explanatory passages can truly provide the justification needed to stave off position skepticism.

For arguments containing counterexamples to work, we some justification for the verdict. As discussed with Gettier, the existence of counterexamples cannot be derived from a second principle. This would prevent the reductio from justifiably negating the target principle, instead merely giving justification that at least one of the two principles is false. Moreover, as argued in Section 2.3.1, in normal non-philosophical contexts justifying counterexamples can be unproblematic – we are simply presented with an object that is a counterexample. In contrast, in philosophical contexts we need some means of grasping

¹¹ See Colaço and Machery (2017) and Machery (2017, 180-181) for a similar two-way distinction.

that an instance is indeed a counterexample, and it is not clear that this is possible without intuitions.

As Deutsch and Cappelen argue, the passages following descriptions of thought experiments may offer a way to supply that justification without intuitions, but the plausibility of this will differ by thought experiment. There is nonetheless at least one general difficulty with this approach. When the statements of the case's intended verdict are taken out, the explanatory passages, on their face, seem to be consistent with the original authors taking the exact opposite position – that the case is not actually a counterexample. Passages after descriptions of thought experiments therefore seem to be doing little justificatory work.

Consider Lehrer's discussion following his Truetemp case in which Mr. Truetemp unknowingly has a machine constantly and accurately feeding the temperature into his thoughts:

Thus, [Mr. Truetemp] thinks and accepts that the temperature is 104 degrees. It is. Does he know that it is? Surely not. He has no idea whether he or his thoughts about the temperature are reliable. What he accepts, that the temperature is 104 degrees, is correct, but he does not know that his thought is correct. His thought that the temperature is 104 degrees is correct information, but he does not know this. Though he records the information because of the operations of the tempucomp, he is ignorant of the facts about the tempucomp and about his temperature telling reliability. Yet, the sort of causal, nomological, statistical, or counterfactual relationships required by externalism may all be present. (1990, 164)¹²

Early in the passage Lehrer states that Mr. Truetemp does not know whether it is 104 degrees. When this is taken out, the passage can be used to make exactly the opposite

¹² This passage has been slightly truncated for space and ease of presentation. The same point being made here can be made about Lehrer's entire passage, albeit less cleanly.

point. Compare Lehrer's original passage to an imaginary version of that passage in which Lehrer's statement that Mr. Truetemp does not know is taken out and one sentence is added to the end:

Thus, [Mr. Truetemp] thinks and accepts that the temperature is 104 degrees. It is. He has no idea whether he or his thoughts about the temperature are reliable. What he accepts, that the temperature is 104 degrees, is correct, but he does not know that his thought is correct. His thought that the temperature is 104 degrees is correct information, but he does not know this. Though he records the information because of the operations of the tempucomp, he is ignorant of the facts about the tempucomp and about his temperature telling reliability. Yet, the sort of causal, nomological, statistical, or counterfactual relationships required by externalism may all be present. *Therefore, Mr. Truetemp can be said to know it is 104 degrees.*

The difference between the two passages comes down to the fact that Lehrer takes himself to have theory-independent access to what knowledge is. For this access to justify his rejection of externalism, it has to come from more than his belief in internalism. Otherwise, because this passage is used to argue against the externalist views of Armstrong, Nozick, Goldman, Dretske, and others, Lehrer would be flagrantly begging the question against his opponents.

This feature is shared by all philosophical uses of imagined counterexamples. To avoid begging the question, we need some sort of theory-independent access to judge whether a counterexample is indeed a counterexample.¹³ Intuitions are exactly the things thought

¹³ Compare this to how some scientific thought experiments seem to work as counterexamples (Gendler, 2000; Kuhn, 1977; Sorensen, 1992). Rather than relying on independent access to the phenomena, some scientific thought experiments seem to reveal contradictions or incoherence inherent in a theory. The thought experiment by Galileo in which a heavier ball is tied to lighter ball and dropped illustrates that in cases such as this, Aristotelian physics predicts that the two balls tied together will fall both faster and slower than the heavier ball alone. Presumably there are philosophical thought experiments that also work by revealing internal contradictions to theories, and these would likely not fall prey to the considerations considered here.

to give us epistemic access to the target phenomena in these cases, such as whether Mr. Truetemp knows the temperature. Unless intuition deniers can explain how we come to justifiably believe things about knowledge (or the topic under consideration in other thought experiments), they cannot provide an account of the justification behind thought experiments sufficient to avoid position skepticism.

2.5.2 Thought Experiments as Positive Evidence and Data Points

Often thought experiments are not used as counterexamples. Instead thought experiments often serve as data points to explain or serve as supporting evidence for a theory. Much the same issues arise here for exegesis-based intuition deniers as with when thought experiments introduce counterexamples.

Thomson's use of variations of trolley cases are a clear example of this method (1976b; 1985). Building from Foot (1967), Thomson treats the trolley cases as points to be explained rather than evidence against one position or another. Utilitarianism is not mentioned. Insofar as the classic presentations of trolley cases by Foot and Thomson serve as counterexamples to utilitarianism, it is because Thomson and Foot use the cases to support the doctrine of double effect. Another candidate of a thought experiment used positively includes Cappelen's (2012, 132-139) reading of Perry (1979). Perry's thought experiments include a professor who goes from believing "the faculty meeting starts at noon" to realizing "the faculty meeting is going on now". According to Cappelen, Perry uses this case and others to point to a previously unstudied phenomena – the change that occurs between the professor believing "the meeting starts at noon" and "the meeting starts now".

As theories successfully manage to explain or predict data points, the thought experiments (or verdicts thereof) serve as positive inductive or abductive evidence for the theory. In such cases, the logical structure of the relevant arguments differs from the

sorts of *reductio* arguments considered so far in this chapter. Therefore, worries about adding extra principles to support the premise do not apply here. Intuition deniers will thus likely have an easier time explaining how such arguments are justified than when the thought experiments are used as counterexamples.

Nonetheless, worries about circularity threaten to prevent exegesis-based intuition deniers from providing an account that provides satisfactory justification in these cases. When considering thought experiments as positive evidence, intuition deniers still need to explain why one verdict is justified over another. Why did Thomson take the trolley cases to be evidence that one ought to pull the lever to save five rather than evidence that one ought to let the trolley kill five? The answer to this question must rely on evidence beyond the theory the case is later used to support, and here the problem for intuition deniers begins to again resemble the problems related to counterexamples. Justification external to the argument is needed, and it is not clear how it can be supplied without intuitions.

2.5.3 Thought Experiments as Mere Illustrations

Intuition deniers will not struggle to explain thought experiments used as mere illustrations of a theory's consequence because the role of these thought experiments is conceptual and rhetorical but not epistemic. Mere illustrations allow authors to communicate a theory by demonstrating the theory's descriptions, predictions, or prescriptions in specific circumstances. These allow authors to pair abstract descriptions of theories with concrete examples. Davidson's Swampman case is one such mere illustration (Davidson, 1987). As discussed by Deutsch, Davidson uses the Swampman case to illustrate externalism about mental content (2015, 117-118). Therefore, the verdicts Davidson draws about the case – that the Swampman would not mean anything he says – is solely a consequence of his theory, not support thereof.

While Davidson argues for externalism about mental content, mere illustrations can be used regardless of whether the author thinks the theory is correct. Nonetheless, using a thought experiment as a mere illustration can provide weak support for the view being illustrated. Being able to coherently describe a scenario using a theory is weak evidence that the view is not incoherent. However, the resulting justification offered in favor of the theory is much weaker than that offered by counterexamples or data points. Moreover, since a verdict is not elicited in these cases over and above what is predicted by the theory under consideration, few metaphilosophers would argue intuitions play a role in such passages.

Intuition deniers have a way to limit the scope of the skepticism they are committed to by showing certain thought experiments are mere illustrations of a theory. It is not clear how many thought experiments this is a plausible understanding of. With exceptions, thought experiments typically occur before any positive proposal in papers – if a positive proposal is even offered. This ordering suggests that thought experiments typically do not serve as mere illustrations.

2.6 Conclusion

In Chapter 1, I introduced the debate between experimental philosophers and exegesis-based intuition deniers. Intuition deniers are motivated by a desire to avoid position skepticism driven by the experimental attack. In this chapter, I argued that in order to avoid the same fate, intuition deniers cannot accept the conclusion that their exegesis reveals unjustified arguments. The thought experiments studied by intuition deniers and experimental philosophers alike have been chosen in part because they have been successful at changing philosophers' beliefs. Thus, if intuition deniers are forced to concede that the authors of thought experiments failed to justify their conclusions, then the subsequent philosophers who have changed their minds in response to the thought experiments have

changed their minds based on bad arguments.

When Gettier's reductio of the JTB analysis of knowledge is formalized, it does not appear as if a satisfactory argument can be produced without appealing to intuitions about knowledge. The argument itself is not formally valid unless Smith's lack of knowledge is understood as being a premise supported externally to the argument, and it is not clear what could support it besides intuitions. Invoking a principle of luck appears to provide the necessary non-intuitive justification, but it problematically weakens Gettier's larger argument against JTB. Moreover, Deutsch does not provide strong reasons to think Gettier's apparent reference to epistemic luck is meant support his verdicts rather than serving as an explanation of the verdicts. Exegesis-based intuition deniers are therefore either committed to the role of intuitions in philosophy or are committed to skepticism about the positions philosophers believe. The former rejects intuition deniers' central empirical claim, and the latter undermines their central motivation.

This chapter and the next two are largely focused on the dissertation's negative thesis. In this chapter I laid out my first line of attack against exegesis-based intuition denial. As presented by Deutsch and Cappelen, it cannot account for the historical influence of some thought experiments without leading to skepticism about the beliefs of people who were convinced by the arguments surrounding the thought experiments. This argument is also, however, the first part of my positive thesis. Historical facts about how philosophy evolved are important to evaluating the epistemology of philosophy. I am going to hold off saying more until the end of Chapter 4, when I can tell a fuller story about which historical facts are important.

In the next chapter, I continue to examine exegesis-based intuition denial, but I take a step back. This chapter showed that responding to the experimental attack via textual exegesis is hopeless. The next two chapters take a step back and argue that the method of textual exegesis is unmotivated from the start. There is no reason for people interested

in the epistemology of philosophy to carefully examine original texts.

Chapter 3

Producer-Focused Metaphilosophy

In the previous chapter, I raised an initial worry about exegesis-based intuition denial. I argued that the project undermines its own central motivation by being committed to skepticism about philosophy. This was an argument focused on the inner workings of exegesis-based intuition denial. I explored what the view itself is committed to about the epistemic status of philosophers based on what they are committed to about the goodness or badness of influential arguments involving thought experiments.

In this and the next chapter, I take a step back and reject the methodological underpinnings of exegesis-based intuition denial. In particular, I discuss the social epistemology of philosophy as a way of exploring the motivation of examining original texts of thought experiments. The project of exegesis-based intuition denial relies on the assumption that a paper's author or a talk's speaker (the work's producer) is more important than the reader or listener (the work's consumer). In this chapter, I reintroduce the distinction between producers and consumers of a philosophical work and highlight why the distinction is important for the exegetical projects considered here. After that, I try to defend the method of exegesis through the social epistemology of philosophy. I find this approach lacking, and in the next chapter I give a positive story for how philosophers learn from each other and explore its consequences.

3.1 Introduction: Exegesis vs Xphi

Experimental philosophers, specifically those belonging to the evaluative or negative experimental tradition, have made inferences about epistemic standing of philosophy and philosophers from psychological studies of people's *reactions* to thought experiments. These designs employed by experimental philosophers focus on the *consumption*, as opposed to the *production*, of thought experiments. Rather than directly revealing something about how thought experiments are created or developed, these experiments investigate how people react to them.

With that in mind, we are in a position to slightly revise the experimental attack, specifically the empirical premise about the psychology of philosophy:

1. According to experimental findings, *consumers'* intuitions are affected by P
2. P is truth-irrelevant
3. As a matter of fact, philosophy involves reconciling intuitions with principles
4. Therefore, philosophy is in bad epistemic standing

The strategy of metaphilosophical textual exegesis rejects the relevance of consumer-focused experimental data, instead taking a decidedly producer-focused approach. They take the perspective of the producer to reject experimental inferences about the epistemic standing of philosophy. Rejecting the relevance of the adjusted step 1 to the conclusion 4, Deutsch and Cappelen present exegetical evidence in original presentations of thought experiments. They argue that intuition-language does not appear in original presentations of influential thought experiments, *and* the texts show no sign that the authors meant to rely on intuitions to establish their conclusions. Thus, insofar as experimental data reveal something about the psychology of philosophy, Deutsch and Cappelen contend these data do not say anything metaphilosophically important. Data on the consumption of thought

experiments, Deutsch and Cappelen argue, are irrelevant in light of evidence about the production of thought experiments.

The question at the heart of this and the next chapter is whether this rejection of consumers is right. As stated, the third step of the experimental attack is ambiguous between producers and consumers – who in philosophy does the reconciling? If we take the producers to be the important drivers of philosophical methods and think that producing a work of philosophy is fundamentally unlike consuming it, then deriving inferences about philosophy from experimental data lacks motivation. Experimental data is looking at the wrong kind of evidence to derive broad conclusions about philosophy. In contrast, if consumers are the most important consideration in philosophical methods, then the method of exegesis is only useful insofar as it reveals historical facts about what authors wrote.

In light of this, we can distinguish between two broad camps of metaphilosophers. CONSUMER-FOCUSED metaphilosophers, namely experimental philosophers and their allies, think that data about the consumption of philosophical thought experiments are key to studying the epistemology of philosophy. PRODUCER-FOCUSED metaphilosophers, in contrast, reject the importance of consumer-focused experimental philosophy in favor of producer-focused exegesis (or in theory, producer-focused experimental philosophy).¹ Notably, neither experimental philosophers nor Deutsch and Cappelen provide robust reasons to think their respective focus on consumers and producers is well-grounded. The metaphilosophical distinction between producer and consumer seems to have only been taken up by Deutsch, and as will be discussed below, his defense of producer-focused metaphilosophy is lacking.

The lack of discussion of consumers vs producers by experimental philosophers is perhaps not surprising. The choice of experimental philosophers to focus on the consumption of thought experiments is probably largely pragmatic. Designing a psychological experi-

¹ These are two extremes. There is of course room for a mixed approach (e.g., Colaço & Machery, 2017).

ment focusing on consumption of predesigned thought experiments allows experimental philosophers to stick closely to familiar and well-established experimental paradigms. It also allows experimenters to limit the range of possible responses to yes/no answers or Likert scales. This leaves the results of studies easy to analyze and fairly straightforward to interpret.

This is not to say that experimental philosophy cannot be producer-focused, but despite the growing sophistication of the methods of experimental philosophers, I am not aware of any experiments that have directly examined the psychological processes behind the creation of philosophical thought experiments. Prior to any actual data on the production of thought experiments, there are reasons to think it is (at least partially) psychologically different than the consumption of thought experiments. Producing a philosophically enlightening fictional story requires clear acts of creativity, while considering such a case does not (Deutsch, 2015, 95-99). This asymmetry is not a foregone conclusion though. Producing and consuming might both involve nearly identical processes of considering cases and eliciting intuitions (see Section 3.2 below). In this case, producing and consuming a thought experiment would differ mainly in that the process of producing the thought experiment is more recursive than consuming it. Nonetheless, it is an open empirical question whether the psychological processes behind creating a thought experiment differ significantly from its consumption. Thus, insofar as experimental philosophy is capable of making epistemic claims about the epistemic standing of philosophers from premises about psychological processes, we should only take it to clearly reveal something about the consumption of philosophical arguments.

Despite the (meta)metaphilosophical break between the competing metaphilosophical strategies of experimental philosophy and textual exegesis, the break has not been discussed elsewhere. Instead, what has gained the most attention is their claim that analyses of the texts of classic thought experiments reveal that thought experiments' authors

argued for their conclusions rather than appealing to intuitions. This strategy has been heavily criticized, including for not properly looking for the sorts of mental states intuitions are (Weinberg, 2014), for failing to explain the historical details of Gettierology (Brown, 2017), for failing to establish the claim that intuitions are not used in philosophy (Egler, 2019), and for being self-undermining (Chapter 2). These arguments all take for granted that the exegetical strategy has something to offer metaphilosophical discourse.

We should nonetheless pay careful attention to the strategy of producer-focused metaphilosophy. Deutsch and Cappelen's producer-focused metaphilosophy has been so influential that it has been adopted by many critics of exegesis-based intuition denial. To illustrate the uptake of producer-focused metaphilosophy in response to Deutsch and Cappelen, consider the following passages in Colaço and Machery (2017), Chudnoff (2017), and Chalmers (2014).

Colaço and Machery (2017) argue against Deutsch's rejection of experimental philosophy's relevance by drawing their own observations about texts. They counter Deutsch's observations about the fact that arguments typically follow thought experiments with their own observations. They retort "the initial presentation of the case is not always, perhaps not typically followed by an explicit argument", using Searle's Chinese Room as an example (Colaço & Machery, 2017, 411).

Similarly, Chudnoff's (2017) criticism of Deutsch and Cappelen accepts their premise of producer-focused metaphilosophy. Disputing Deutsch and Cappelen's argument-based picture of thought experiments, Chudnoff appeals to how Lehrer (1990) discusses the Mr. Truetemp case, saying

Lehrer does not just describe Mr. Truetemp, state his verdict about the case, and leave it at that. He says some additional stuff. The stuff doesn't constitute a good argument for his verdict however. Rather, it helps us see what he sees about Mr. Truetemp. (Chudnoff, 2017, 383)

Chudnoff argues for intuition-based methodology based on the observation that Lehrer himself intends to invoke in the reader intuitions against externalist theories of knowledge,² thereby adopting producer-focused metaphilosophy too.

Chalmers (2014) also adopts producer-focused metaphilosophy in criticizing Cappelen. Discussing the role of intuitions in his own work on philosophical zombies, Chalmers says,

In that presentation, I first appeal to the conceivability of zombies, saying I take this to be intuitively obvious (and therefore noninferentially dialectically justified), but I go on to give a number of arguments for this claim from underlying principles (for example, the nonanalyzability of consciousness in functional terms). (2014, 540)

In this passage, Chalmers is doing producer-focused metaphilosophy on his own work. Because Chalmers is the producer of the work being analyzed, Chalmers can beat Cappelen at Cappelen's own game. While Cappelen analyzes the methods of Chalmers (1996) (the text) through reading, only Chalmers (the producer) has first-hand knowledge of the methods of Chalmers (1996) (the text). Chalmers takes *his own methods* to ultimately defend the use of intuitions in philosophy.

3.1.1 Outline

This chapter and the next examines whether the exegetical strategy is methodologically justified in the first place. Specifically, I answer the following question:

If we are interested in the epistemic standing of philosophers, who should we focus on? The methods, psychology, and/or justification of the *producers* of thought experiments, the methods, psychology, and/or justification of the *consumers* of thought experiments, or both?

² I argue for a similar but more general picture of the epistemology in Section 4.1.

In order to answer the question, in this chapter I identify the most plausible defense of producer-focused metaphilosophy, namely philosophical testimony. If reading a text produces testimonial philosophical knowledge, then, at least on certain anti-reductionist accounts of testimony, our epistemic standing qua philosophers depends on the epistemic standing of the original producers of philosophical arguments.³

In the next chapter, I give my positive account of how philosophers do learn from each other. While I argue that we have testimonial knowledge of some philosophical theses, this testimony comes from sources other than the producer of a work of philosophy. Instead, knowledge that is caused by a work of philosophy is based on the epistemic tools of *consumers* who consider texts by their own lights. Therefore it is the consumers, not producers, that carry the most importance for the purposes of evaluating the epistemology of philosophy, and exegetical analyses of original texts carries very limited relevance to such a project.

As a way of introducing the dialectic, in Section 3.2 I introduce the first discussion of philosophical producers vs philosophical consumers that I am aware of: Deutsch's discussion of the process Gettier must have gone through in producing Gettier cases. I argue that this discussion does not do the work Deutsch wants or needs it to. In Section 3.3, I then discuss other explicit defenses of producer-focused metaphilosophical methods. After criticizing these, Section 3.4 introduces philosophical testimony as producer-focused metaphilosophy's best hope. Section 3.5 argues that peer disagreement defeats a significant amount of philosophical testimony, raising a significant challenge to producer-focused metaphilosophy.

³ When talking about "philosophical knowledge" and "philosophical testimony", I specifically mean knowledge and testimony about the target of philosophical inquiry. These are substantive, object-level claims about the sorts of things philosophers investigate as philosophers. Compare this with claims about philosophy as a discipline. "Berkeley's idealism is false" is an example of the sort of proposition whose epistemology I am interested in investigating, while "Berkeley's idealism is widely considered to be a major development of British empiricism" – a claim about the field of philosophy – is not. I am assuming for present purposes that we have such knowledge. Starting in Chapter 5, I am neutral on the point.

3.2 Focusing on Producers

As mentioned, the distinction between producers and consumers is originally Deutsch's. One of the explicit mentions is in relationship to the order of explanation objection (see Section 2.4). Deutsch raises the distinction as one of his main arguments against the reading of "mere coincidence" as a purely abductive inference meant to explain why Smith does not know. Instead, Deutsch asks us to consider what happened to Gettier when he developed the thought experiments in the first place:

If intuitions are evidence for [the verdicts], they are not Gettier's intuitions, for the simple reason that Gettier almost certainly did not come to believe [the verdicts] on the basis of intuition. Arriving at that belief, for Gettier, did not involve the passive sort of cognizing characteristic of intuiting that something is so. It must instead have involved active "thinking through" the details of the cases, and reasoning that his conditions about them were true. [emphasis original] (Deutsch, 2016, 85)

Deutsch's point about the asymmetry between the cognitive aspect of producing a thought experiment and consuming it is correct and important. When we discover counterexamples, they do not simply pop into the mind out of nowhere (Sorensen, 1992). Creators of thought experiments, whether they trade in intuitions or arguments, are not operating in total darkness when trying to develop a counterexample. There are many possible theories that conflict with the claim that JTB is necessary and sufficient for knowledge, but only some of those theories are plausible. More importantly, we have some sense of which conflicting theories are the plausible theories. Someone trying to produce a counterexample to an analysis of knowledge knows not to look for a counterexample based on the principles *beliefs formed on Tuesdays cannot be knowledge* or *we cannot know anything about cats* – despite these principles' incompatibility with the

JTB account.

A successful description of how a thought experiment is created will therefore explain why Gettier was able to hone in on a counterexample rather than blindly searching for one. Elsewhere, Deutsch tries to explain this by using the analogy of a recipe (2015, 98). Having justification to think epistemic luck precludes knowledge even if the belief is justified and true, Gettier “cooked” up a case with a counterexample knowing prior to any intuitions what was going to come out of the oven. Deutsch is arguing that because this creative process is undoubtedly active and requires careful and intense thought, it must not require intuitions. Instead, the verdicts must have been inferred.

Here Deutsch’s argument makes a mistake common to the intuition literature, especially literature that talks about the relationship between intuitions and reasoning. Careful, effortful, and deliberate thought is not inconsistent with the role of intuitions.⁴ Rather, reasoning and intuition (insofar as they are distinct) should be seen as tightly integrated cognitive strategies that recursively interact (Kornblith, 2012; Nagel, 2012a). Reflection guides which intuitions are formed, and those intuitions then guide the direction of our reflection. The recursive nature of intuitions and reflection should be well-known to anyone who, as an audience member in a philosophy talk, finds their understanding of something the speaker said unintuitive. The charitable thing to do would be to find alternative understandings of what the speaker is defending (a reflective task) in order to check whether those too are unintuitive. This process may repeat itself through multiple iterations as we adjust our attempts to interpret our understanding of the speaker until we find something we find intuitively plausible.⁵

Because of the back and forth relationship between intuitions and reasoning, we can, contrary to Deutsch’s argument, tell a story of the production of thought experiments that is compatible with both a reasoned production and intuition-based justification:

⁴ See (Cappelen, 2014a, §2.4) for another example of the mistake.

⁵ Admittedly this example is not theoretically neutral on the nature of intuitions and intuitiveness. See Chapter 7 for further discussion on the point of theoretical neutrality here.

One day a man named Edmund was sitting around and decided he needed to publish a paper to secure tenure. Edmund suspected the JTB analysis of knowledge was flawed. He also suspected knowledge was incompatible with luckily true beliefs. To test this, he wrote down a story where a character has justified true belief but the belief was true by mere coincidence. While writing this story, he took great care to consider and remove possible confounds. Edmund then asked himself, “Is this a counterexample to the JTB analysis?”. Considering the case carefully, he intuited that Smith does not in fact have knowledge, leading Edmund to exclaim “Ah, it is true, for I have intuited it!”.⁶ Edmund then writes a brief introduction introducing the JTB analysis, writes a second story that also causes him to intuit that Smith has JTB that is not knowledge, and writes a conclusion.

Given the coherence of this imagined producer-focused origin story behind the production of Gettier’s paper, Deutsch’s picture of creation does not preclude intuitions from playing a key role in the creation of thought experiments. Indeed, it shows that the distinction between producer and consumer might be quite muddled, since in the thought experiment Gettier is in a sense consuming the thought experiment as well as producing it. On any account of intuition-based production of thought experiments, the producer needs to go through a creative process, and that process may involve following a recipe. While Deutsch’s picture implies that the chef in the analogy knows that what he is making will turn out, on an intuition-friendly picture, the chef might not. Intuitions behave in such a way that we cannot always predict their outcome. Instead, at least on an intuition-friendly understanding of metaphilosophy, everyone, including producers, need to check intuitions in order to know what a thought experiment shows.

Thus, focusing on the production of a thought experiment so removed from the actual

⁶ This exclamation, while perhaps not necessary for the thought experiment, makes it clear that intuitions are playing a role in the state sense (see Section 1.4.1).

process shows nothing itself about the role of intuitions in philosophy. Nonetheless, this perspective on thought experiments is central to the exegetical project, and an important shift from the consumer-focused discussion found among experimental philosophers. Reading into the text, according to producer-focused metaphilosophers, is important for metaphilosophy, although we cannot know how it is important to metaphilosophy without checking the text.

3.3 Existing Defenses

In this section, I examine and reject the defenses of producer-focused metaphilosophy given by producer-focused metaphilosophers. The closest thing to an initial articulation of producer-focused metaphilosophy is the following passage by Deutsch:

The important methodological question is: What methods are employed by *good* philosophers, ones who, by fairly wide consensus, have made interesting and important philosophical progress, increasing, in a significant way, our body of philosophical knowledge? (2015, 41):

In other words, there is something particularly important about paying attention to original and groundbreaking works of philosophy. The original texts are more metaphilosophically relevant than anything “downstream”, such as how they are understood or how they are discussed by other philosophers.

Cappelen implicitly takes a similar position. In criticizing the negative experimental philosophy of Swain et al. (2008), which studied intuitions about Lehrer’s Mr. Truetemp cases. Notice Cappelen’s focus on Lehrer himself, especially in Cappelen’s first point:

Swain et al. make two false assumptions about Lehrer’s paper: [First,] *Lehrer’s* judgment that Mr. Truetemp does not know is based on an intuition. [Second,] The failure to account for that intuition *is used* as an argument against

reliabilism. [emphasis added] (2012, 222-223)

Cappelen is interested in the epistemic status and methods of Lehrer himself, as presented in Lehrer (1990). In the quoted passage, Cappelen argues Swain et al. did not study the actual methods of Lehrer's paper. Instead, he claims Swain et al. make a false assumption about Lehrer's methods in their experimental design and resulting criticism of the case method. Therefore, Swain et al.'s discovery of order effects in judgements about Mr. Truetemp cases does not matter to the debate about reliabilism, Cappelen argues, because *Lehrer* does not rely on such judgements as evidence.

Further justification is needed for Deutsch and Cappelen to secure their rejection of the experimental attack. Suppose contrary to the previous chapter that original presentations of thought experiments do present non-intuitive evidence for the verdicts of thought experiments. It is a further claim that evaluation of this non-intuitive evidence is important for evaluating the methods of philosophy because it may be that the actual contents of a text are nearly or completely epiphenomenal to later practice. The very same ignorance of the content of original texts that Deutsch and Cappelen lament offer credibility to this possibility. The conclusions of many influential analytic texts are often discussed and taught without actually being read (much to both Deutsch's and Cappelen's frustration). For example, people who have never read *Naming and Necessity* believe that Kripke disproved descriptivist theories of reference using a series of thought experiments involving mistaken definite descriptions because this is how the main thrust of Book 2 of *Naming and Necessity* is presented in classrooms and at the pub. These later presentations and discussions of texts may be the source of knowledge and justification for most or all philosophers.

Producer-focused metaphilosophy denies the relevance of these later presentations and later consumption of the material, thus justifying the move Deutsch and Cappelen both make from text-based evidence to a claim about philosophical practice more generally.

As far as Deutsch provides a defense for producer-focused metaphilosophy, it seems to be the following claim:

The core methods of the discipline, and of any discipline, are reflected most clearly by the most clearly successful examples of discovery and progress in the discipline. A focus on such examples in philosophy is entirely appropriate. (2015, 41)

This claim about methods and progress is simply not true. The methods of disciplines are not reliably reflected in the most notable examples of progress because progress has often occurred when core methods are tweaked or set aside. Darwin’s “On the Origin of Species” (2011) relies on an argument from analogy between domesticated animals and wild animals. Arguments from analogy are rarely used in contemporary biology. Turning to thought experiments, thought experiments are fairly uncommon in primary scientific literature. Nonetheless, some of the greatest advances in physics – whether the shift to Newtonian mechanics, special and general relativity, or quantum mechanics – involved thought experiments (Gendler, 2000; Kuhn, 1977). Similarly, there are many norms of philosophy *not* reflected in Gettier’s paper. Few other published philosophy papers have successfully defended a thesis in 1000 words using thought experiments.

Setting aside this worry about whether extraordinary works of a discipline use ordinary methods, the rate at which philosophers misunderstand each other raises a second objection to producer-focused metaphilosophy. Philosophers misread, mishear, and misunderstand works of philosophy all the time. Many well-known and generally accepted works of philosophy are subjected to widespread disagreement about the works’ content or intentions. Philosophers cannot agree on something as basic as the content of Gettier intuitions,⁷ or even if Gettier used intuitions in the first place. Despite this, philosophers generally agree that JTB is not knowledge. Producer-focused metaphilosophers must then

⁷ See Chudnoff (2011) §7 and Cappelen (2012) Chapter 3.

explain how even simple texts like Gettier (1963) can be widely misunderstood but still be more metaphilosophically relevant than the epistemic states of the people reading them.

Third, and relatedly, if intuition deniers are right and philosophers have mistakenly thought that intuitions are used as evidence in philosophy, then philosophers have been widely mistaken about the methods of clearly successful examples of progress. This leaves Deutsch in an awkward situation of accepting that progress has occurred but philosophers have been widely mistaken about what the progress is grounded in or even consists of. Some philosophical accounts of progress might allow for this,⁸ but even if we accept that progress can occur in philosophy when a substantial majority of philosophers have false beliefs about what the progress consists in, we still have reasons to not take original texts very seriously. Assume that authors of notable thought experiments did not rely on intuitions but successfully discovered philosophical truths. Assume further that consumers reading the text thought that intuitions were involved and treated their intuitions as putative warrant for the conclusions being defended. Then metaphilosophers still have to accept that many philosophers' beliefs have been caused by unreliable methods, leading them to defend false beliefs and to explore blind alleys (see Nado, 2016).

Existing defenses of the method of exegesis employed by Deutsch, Cappelen, and many subsequent critics fall short. Deutsch appears to have provided the only substantial defense of the method. His defense, that notable works of philosophy are illustrative of the methods of the rest of philosophy, faces serious and likely insurmountable challenges. We therefore need to look elsewhere to try to defend producer-focused metaphilosophy.

⁸ Bird (2007) is one possible account. Appealing to Deutsch's use of "progress" in the passage quoted above is one possible route for defending producer-focused metaphilosophy that I leave unexplored here. It does, however, face two immediate hurdles. First, evaluative metaphilosophy has traditionally focused on warrant, justification, or knowledge of philosophical propositions instead of other epistemic states such as progress or acceptance. To defend producer-focused metaphilosophy this way we need reason to shift the focus of evaluative metaphilosophy to these other epistemic states. Second, such a defense of producer-focused metaphilosophy still needs some way of screening off later misunderstandings as metaphilosophically unimportant. Progress is likely not much good to a community of inquirers if most or all of the inquirers misunderstand the progress, such as mistakenly thinking that intuitions are involved in the discovery.

3.4 Testimony and Philosophy

For the rest of this chapter, I explore the potential of philosophical testimony, specifically anti-reductionist accounts of philosophical testimony, as a way of defending producer-focused metaphilosophy. I take this to be the most plausible defense because it, at least in theory, allows for a robust epistemic connection between producers and consumers. If this connection is in place, producer-focused metaphilosophers have a ready reason to look at original texts. If we gain knowledge from producers of philosophical arguments via testimony, then learning about the sources of justification of original authors reveals something about the justification of those who read the text.⁹

Naturally, accounts of testimony differ on what exactly is necessary for testimonial knowledge or justified testimonial beliefs, and the differences between accounts matter to the sort of defense philosophical testimony provides for producer-focused metaphilosophy. The main faultline in the testimony debate turns on whether testimonial warrant, justification, or knowledge depends on non-testimonial evidence. REDUCTIONISTS argue that the uptake of testimony depends on the listener drawing upon other sources of evidence, such as perceptual evidence or inductive inferences (Coady, 1992; Fricker, 2014; Hume, 1748; Van Cleve, 2006). In contrast, ANTI-REDUCTIONISTS deny the need for additional inferences or evidence from non-testimonial sources, but they typically grant that these non-testimonial sources can nonetheless provide defeaters for a given piece of testimony (Audi, 1997; Burge, 1998; Reid, 1769).¹⁰

To flush out the distinction between the two camps, reductionists argue that testimonial justification depends on other sources of justification. To have testimonial knowledge the listener must have positive reasons for believing what the speaker says. On such a

⁹ Given that testimonial chains seem to be possible – where one testifies what was testified to them – the examining original text could tell us something about the justification of most or all people downstream as well.

¹⁰ See Lackey (2008) for an influential position that blends both camps. See also Kusch (2002) and Hawley (2010). Thanks to Katharina Bernhard for invaluable help separating out these different camps.

reductionist account, justification for testimony comes from sources deeper down, most often foundational sources of knowledge like memory, perception, or inference (Kusch, 2002, 29). Hume (1748), often cited as the earliest reductionist, argues

The reason why we place any credit in witnesses and historians, is not derived from any *connexion*, which we perceive *a priori*, between testimony and reality, but because we are accustomed to find a conformity between them.

We trust testimony, according to Hume, because experience has taught us that testimony about known facts has generally been reliable, and so we expand our trust to testimony about unknown facts. More recently, Van Cleve (2006, 69) argues “testimony gives us justified belief and reflective knowledge not because it shines by its own light, but because it has often enough been revealed true by our other lights.” Other sources of knowledge have corroborated testimony, allowing us to gain justification from testimony. Testimonial knowledge on the reductionist view therefore requires the speaker to have positive reasons in favor of the testimony, such as inductive evidence that the sort of testimony we have is generally reliable.

In contrast, anti-reductionists about testimony argue that testimonial justification is basic and *does not* depend on any further source of justification. Audi (1997, 408) draws an analogy between testimony and perception, saying that like perception, testimony is “constrained by other beliefs without being inferentially based on them”. While beliefs affect what we believe from testimony, testimonial knowledge is not inferred from those beliefs. Because anti-reductionists hold that testimony is in some sense basic, no positive reasons are needed to justify testimony. Just like we do not need positive reasons to be justified in our perceptual beliefs but can lose justification if we find out we are hallucinating or something else has gone wrong, testimony on anti-reductionist accounts remains justified unless we have reasons to doubt it. That is, testimony has default justification unless defeated.

I will be assuming anti-reductionism for the purposes of criticizing producer-focused metaphilosophy for two reasons. First, because reductionist accounts typically require positive evidence for the testimony in addition to the requirement shared by anti-reductionists that there are not any defeaters, it is easier to establish on an anti-reductionist account that a belief is testimonial knowledge. This makes it easier for producer-focused metaphilosophers to cross the hurdles necessary to say that we have testimonial knowledge of the contents of original works of philosophy.

Second, anti-reductionism bolsters the motivation behind producer-focused metaphilosophy by strengthening the epistemic relationship between producer and consumer. In particular, many anti-reductionist accounts readily allow for transmission of epistemic states from producer of testimony to consumer of the testimony (Greco, 2016). When this sort of transmission occurs, a hearer knows a proposition on the basis of the speaker's own justification and not, say, that the speaker testified. Assuming transmission, Gettier's own reasoning and arguments are the reasoning and arguments handed down to us via testimony. They are, in some sense, our reasoning and arguments as well. Thus, we have a clear reason for saying why the methods of influential authors matter to the epistemology of philosophy.

3.4.1 Do Philosophers Testify?

Speaking anecdotally, philosophers are often wary about the idea that they gain philosophical knowledge through testimony. In the same way testimonial knowledge of an artwork's beauty seems problematic, there is a vague sense that philosophers' reliance on testimony would constitute a professional failing or a failure of intellectual virtue. Nonetheless, there are two main reasons to think that testimony plays a key role in our consumption of other people's philosophy.

First, a substantial amount of our knowledge in other domains depends on the knowl-

edge of other people. I have never been to Spain, but my knowledge that Barcelona is sunny in the summer and cloudy in the winter depends (at least in large part) on what my Spanish and Catalan friends have told me. Closer to philosophy, logicians and mathematicians seem to have testimonial knowledge of *a priori* and necessary facts based on the work of one another. Unless we have reasons to think that philosophy is different than other domains, denying philosophical testimony risks entailing the implausible claim that testimony in areas related to philosophy such as mathematics and logic is impossible (Ranalli, 2020).

Second, philosophers *act* as if we gain philosophical knowledge through testimony. While work in philosophy is piecemeal, it is nonetheless interrelated. Advances in one topic affect advances in another, and our philosophical reasoning has assumptions based on works of other philosophers that we lack relevant expertise, time, or desire to evaluate. The most charitable reading of this behavior is that we are taking our beliefs as being justified by the testimony of experts in other sub-disciplines. Which advances we treat as if we know through testimony will differ from philosopher to philosopher, but potential examples include that possible worlds are the right way to handle modal reasoning and modal semantics, that the law of non-contradiction holds, or that ZFC set theory is most likely self-consistent. If it turns out that we do not know these propositions through testimony despite our use of them in our philosophizing, we have a devastating result for the epistemic standing of philosophers.

Note, however, that assuming the work of others in our own work has at least one other reading. We may be drawing conditional inferences from what base claims seem best or most likely to be true.¹¹ Because this behavior is widespread, this would mean that most philosophical knowledge is conditional as opposed to non-conditional, but some philosophers – especially those with Quinean and/or anti-realist leanings – might find this palatable. While removing the need to justify the antecedents of conditionals via

¹¹ Thanks to Mark Phelan for suggesting this possibility.

testimony, this does not remove all potential need for testimony in philosophy. Even if philosophical knowledge is mostly conditional, we may also lack relevant expertise, time, or desire to determine whether certain propositions follow from our assumptions.

Philosophical testimonial knowledge therefore seems to square with philosophers' behavior and our sense that testimony is possible in closely related *a priori* domains. This only goes so far as establishing that philosophical testimony fulfills theoretical and practical desiderata. Determining whether we actually have such knowledge is another task altogether.

One obvious way forward is to investigate whether consumers of a work have linguistic understanding of the utterance and believe the assertions made by the work's producer. Linguistic understanding and belief are not jointly sufficient for testimonial knowledge though because defeaters can prevent testimonial knowledge. If I have an undefeated defeater for the speaker's testimony, I can understand and believe what they assert without knowing what is asserted. Therefore, in order to determine whether assertions produce testimonial knowledge, we need to examine the features of philosophical communication in order to determine whether you know the content of this sentence through testimony.

3.4.2 Impure Testimony

Besides defeaters, there is another wrinkle to examining the extent of testimonial knowledge in philosophy. As Anscombe (1979, 145) observed (see also Wanderer, 2013), the epistemology of statements in philosophy differs from the epistemology of statements in many other areas.¹² Philosophers communicate to each other through arguments, and this is important for understanding how philosophical consumers learn from the works of philosophical producers.

To see the wrinkle, consider this toy example of an (imperfect) anti-skeptical argu-

¹² This is also implied by Moran (2006, 279-280).

ment:

I know that I have a hand. My hand is external to my mind, so if I know I have a hand, I know there is an object external to myself. Given closure of known entailments, if I know there is an object external to myself, I know there is an external world. Therefore I know that there is an external world.

The toy argument does not map nicely onto the examples typically given of testimony by philosophers. Compare the argument to examples from the testimony literature, such as being told it is cold outside (Moran, 2006, 278), reading first-hand accounts of religious miracles (Hume, 1748), receiving directions to the post office (Coady, 1992, 38), and a self-proclaimed clairvoyant telling a friend that Elvis is alive in San Diego (Lackey, 2008, 16). In these cases we might have reasons to trust or distrust what is asserted, but this falls short of our relation to the premises and conclusions of the toy argument.

The toy argument is consumer-evaluable in a way canonical examples of testimony are not. The sorts of propositions communicated in “traditional” cases of testimony are propositions that we must go out of our way to check for ourselves, if they are even within our epistemic grasp at all. In contrast, when reading the anti-skeptical argument above we are able to evaluate – on the fly – the claims being defended. I can evaluate for myself whether the premises entail that I know there is an external world, whether the argument begs the question, or whether closure of known entailment holds.

The distinction between the toy argument and the cases from the testimony literature is not as clear cut as it might first appear, however. How much they resemble each other depends on how the details are filled in. Consider Coady’s example of asking directions to the post office. I may not know the town at all and just urgently need to buy some stamps while on vacation. I do not know the town so when I get directions I have to take the testifier’s word for it. As far as I know how to find the post office, I know because of her testimony. Imagine instead that I have been in town a few days and just want

directions to double check my route. In this second case, I have the ability to evaluate the directions I receive, but I still gain justification from them.¹³

This speaks to an epistemic feature of testimonial knowledge. Sometimes testimony is the only justification for a proposition, such as, plausibly, my knowledge of the propositions *that everything is made up of atoms, that Tashkent is the capital of Uzbekistan, and that Obama's middle name is Hussein* (see, however Anscombe, 1979). In contrast, sometimes testimony is merely one source among many. The COVID-19 Risk Assessment for my office building tells me my office building is actually three buildings connected by hallways. This is a fact I had already suspected based on stray observations about the locations of stairwells, differences in exterior stone-work, the way plumbing seems to be connected, etc.

To help keep the difference straight, we can call instances where testimony is the only justification for a proposition PURE TESTIMONY.¹⁴ Pure testimony occurs when the only source of justification for the communicated proposition is based on the communication act, such as the actual or perceived justification of the producer. Contrast this to IMPURE TESTIMONY where testimony is merely one source of justification among others. With impure testimony, testimony provides a distinct but complementary epistemic basis for knowledge.¹⁵ Notice that testimony that is principle impure need not be impure in

¹³ Notice that there is a risk of double-counting evidence where me and the testifier both have the same evidence for p and I nonetheless take their testimony as additional evidence for p. Thanks to Jan Constantin for this point.

¹⁴ Thanks to Jessica Brown for this distinction.

¹⁵ For a similar distinction, see the distinction between moral advice and moral testimony in Hills (2009) and Sliwa (2012). I do not take myself to be interacting with the aesthetic or moral testimony debates for two reasons. First, what is true in ethics and aesthetics might not be true in areas like epistemology and philosophy of language (Ranalli, 2020). Second, the moral and aesthetic testimony debates are often interested in every day cases of moral and aesthetic testimony, such as asking a friend for a movie recommendations. In this paper, I am instead interested in how philosophers communicate as part of the practice of philosophy. Also note that the pure/impure testimony distinction is itself distinct from the difference between reductionist and anti-reductionist accounts of testimony. The reductionism/anti-reductionism debate is about where the justification for testimony comes from – whether it is basic or derivative on something like perception. The pure/impure distinction is about whether justification for the proposition testified comes from solely testimony or if it comes from testimony and something else.

practice. I might have justified beliefs about the best way to the post office but forget my justification, and even though I have the right justification to treat the testimony as impure, I fail to bring my own justification to bear on the issue.

3.4.3 Challenging Producer-Focused Metaphilosophy

This chapter has thus far explored producer-focused metaphilosophy and tried to find justification for the method of examining original texts for the purposes of studying the epistemology of philosophy. I have argued that the best strategy for producer-focused metaphilosophers, such as exegesis-based intuition deniers, is to make certain claims about the epistemology of philosophy. Specifically, I have argued that testimony creates the sort of strong epistemic link between producer and consumer needed to motivate producer-focused metaphilosophy. As the last few pages have demonstrated, showing that philosophical beliefs are testimonial knowledge is far from trivial. First, defeaters may prevent the beliefs from being knowledge. Second, the use of other epistemic sources may prevent knowledge from being testimonial.

Motivating producer-focused metaphilosophy by appealing to the social epistemology of philosophy, however, requires any motivation to rely on an adequate story of the social epistemology of philosophy. It is no use defending something on the grounds of the social epistemology of philosophy if it relies on an unrealistic or problematic account of how philosophers learn from each other. There are two ways to flush out this challenge. First, producer-focused metaphilosophers need to capture the actual practice of philosophers (call this the DESCRIPTIVE ADEQUACY CHALLENGE). They need to defend producer-focused metaphilosophy based on the actual social epistemology of philosophy. Second, since I am assuming a general optimism about our ability to know things in philosophy, defending producer-focused metaphilosophy also needs to explain how it is that philosophers come to know what they know. Call this the EPISTEMIC ADEQUACY CHALLENGE.

Now that I have done the stage setting required, I can argue that they will struggle to meet either adequacy challenge.

The next chapter is devoted to the descriptive adequacy challenge. I propose an account of how philosophers learn from each other that involves almost no testimony. Indeed, rather than just being a problem for producer-focused metaphilosophers, my account lends support to the opposing view, consumer-focused metaphilosophy.

In the remaining section of *this* chapter, I raise a challenge for the epistemic adequacy of producer-focused metaphilosophy. As I have discussed in this chapter, testimony is the best shot producer-focused metaphilosophy has for motivating the view. However, very few acts of communication in philosophy are capable of producing philosophical testimonial knowledge because peer disagreement defeats any putative warrant they might have.

3.5 Peer Disagreement and Philosophical Testimony

Philosophical arguments, on their face, look like impure testimony. I am not locked in to taking an author at their word. When I read Putnam's defense of semantic externalism, for example, I have the chance to evaluate the arguments and check using my own tools to see if Putnam's arguments are good arguments for his conclusion. At the same time, Putnam's arguments are written in the form of assertions and it seems I have rational pressure to think semantic externalism is true regardless of my evaluation of Putnam's arguments. After all, Putnam was a much better philosopher of language than I am (van Inwagen, 2010).

This picture of philosophy as impure testimony is too rosy. While philosophical communication might in principle produce both testimonial and non-testimonial knowledge, there are serious reasons to doubt that such communication produces testimonial knowledge in practice. Specifically, the extent of peer disagreement looks to defeat most poten-

tial testimonial knowledge and creates a serious challenge for producer-focused metaphilosophy.

3.5.1 How Widespread is Disagreement?

The extent of peer disagreement in philosophy is an open empirical question (more on this in Section 4.6). Some philosophers writing on peer disagreement in philosophy have taken cues from the PhilPapers survey (Bourget & Chalmers, 2014). The survey collected data on professional philosophers' verdicts about thought experiments and other theoretical commitments. Interestingly, philosophers writing on the survey have come away from the survey with opposite conclusions. Both Grundmann (2013) and Chalmers (2015), for example, argue that the results show evidence of widespread disagreement and lack of convergence in opinion among philosophers. In contrast, Ranalli (2020) points out that for many questions, agreement among experts on the topic is fairly high. For instance, 81.6% of specialists about skepticism responded that we can know things about the external world (Kelly (2005) makes a similar observation without survey data).

It is beyond the scope of this dissertation to give a robust quantitative or qualitative analysis of the PhilPapers results, but apparent agreement in the survey hides deeper disagreement. Even if we accept Ranalli's observation that experts on topics like skepticism or Berkeleian idealism generally agree on big-picture questions, surface-level consensus hides deeper disagreement. Even where philosophers broadly agree about philosophical questions, they tend to disagree about the details. While philosophers working on skepticism generally agree that we know things about the external world, they disagree about whether we know things about the external world because of content externalism, abductive arguments, Moorean arguments, closure failure, or other reasons. Often then a philosopher may believe a position P on grounds X , another philosopher believes P on

grounds Y, and X and Y are incompatible.¹⁶ Therefore, even when philosophers generally agree on the truth of a proposition, there is often still peer disagreement about that proposition.

Setting aside the PhilPapers survey, even without quantitative analysis it is clear that in philosophy there is quite a lot of peer disagreement. Like other academic disciplines this disagreement is at its most widespread and salient at the cutting edge of the field.¹⁷ Disagreement at the cutting edge is just a feature of novel work in any field. Because cutting-edge work tackles issues we do not fully understand, it is messy, difficult, and involves false starts.

3.5.2 Disagreement Matters

In Section 3.4 I argued that producer-focused metaphilosophers have an easier time on anti-reductionist accounts of testimonial knowledge. It is worth noting however that when it comes to peer disagreement, the reductionist and anti-reductionist distinction is largely orthogonal to the problem peer disagreement raises for producer-focused metaphilosophers. For testimony to be insufficient for knowledge on both sorts of accounts, there need to be negative evidence to distrust the testimony, although reductionists will also say knowledge is not produced by testimony if there is an absence of positive evidence to believe the testimony. Extensive and persistent peer disagreement defeats testimony in both ways. It provides negative evidence for taking the word of a philosopher as well

¹⁶ The wide range of bases for generally accepted philosophical positions might be a further reason to worry about the testimony of philosophy because it suggests (but falls far short of definitively demonstrating) that philosophers are more interested in finding good enough ways to defend predetermined positions than the right positions.

¹⁷ An important and large exception is in many fields of humanities other than philosophy where disagreement, while widespread, appears to be less focal to work in the field than the disagreement is in philosophy. Marxist and feminist historians, while drawing far different conclusions about a given historical event, may take their work as merely two different ways of analyzing the event as opposed to being mutually exclusive in the same way philosophers take opposing views to be mutually exclusive. I suspect attitudes like this are due to the greater influence of anti-realist philosophers in fields of the humanities that are not analytic philosophy.

as preventing the consumer from having positive evidence for believing them.

Looking at the lack of positive evidence first, it would be arbitrary for us to take the side of one of the peers engaged in disagreement over the other (Ranalli, 2020; Sliwa, 2012). This is especially true if we lack the expertise or motivation to check the soundness of the argument or the truth of the proposition by ourselves.¹⁸ As I discussed in Section 3.4.1, philosophers do not always read works of philosophy with a critical eye. This includes times where we do not know the background considerations and assumptions of a debate or subdiscipline well enough to trust our ability to judge the soundness of an argument. When we are in this position of relying on a speaker for our justification, forming beliefs based on the assertions of one philosopher among disagreeing peers would be arbitrary.

Besides preventing us from having positive reasons to believe the testifier, peer disagreement also defeats testimony. Depending on the features of disagreement, disagreement can be evidence of a lack of progress or knowledge among the peers. Kornblith (2010, 2013) argues, for example, that the pervasive long-standing disagreement in philosophy is evidence that we lack philosophical knowledge. Kornblith argues that unlike in science, where debates tend to last a few decades, the general outlines of many of the more longstanding debates in philosophy have been in place as far back as the ancient world. At least in western philosophy, today's debates in the philosophy of mind between materialists and dualists and in metaphysics between naturalists and non-naturalists existed in some form between Democritus and Plato (Robinson, 1995, 60), and the main sides in contemporary debates on mereology date back to Democritus and Anaxagoras (Ladyman, Ross, Spurrett, & Collier, 2007, 20). These sorts of long-standing unresolved

¹⁸ Following the PhilPapers study and Ranalli (2020), I am using “expert” and “expertise” in a neutral way to describe people who work professionally on the topic, not people who are, say, more reliable than non-experts at coming to the truth in their area of expertise, e.g., Constantin and Grundmann (2018). I am talking about experts for ease of presentation, and what most of what I say about experts below could be replaced with “people who have had a good think about the topic”, for which we use as a proxy that they work professionally in the relevant subfield. Note that this ambiguity of “expert” has been noticed elsewhere, e.g., Ball (2020); Sawyer (2018).

debates in philosophy, Kornblith argues, are evidence that the methods philosophers use are incapable of producing knowledge, at least in these debates. Therefore, peer disagreement is a defeater for all party's testimony because we have evidence that the peers engaged in the debate are not good at finding the truth.

Peer disagreement in philosophy therefore looks potentially devastating for producer-focused metaphilosophy, but there are ways producer-focused metaphilosophers could push back against this worry in order to meet the epistemic adequacy challenge. They could, for example, argue that knowledge is not really the goal here. Rather philosophers aim for something much more epistemically lightweight, such as acceptance, that does not involve commitment to the truth of what is being asserted (Everett, 2014; Plakias, 2019). Fully fleshing out this response is beyond the scope of the present discussion. Following the literature I am discussing, I am interested in the epistemically heavyweight states of justified philosophical belief and philosophical knowledge. Arguing philosophy aims for something like acceptance is a huge dialectical shift and one producer-focused metaphilosophers would need to defend.¹⁹

Instead, producer-focused metaphilosophers' most obvious move to meet the epistemic adequacy challenge is to argue that peer disagreement in philosophy is not as pervasive as I have suggested. While disagreement is pervasive and longstanding in some areas of philosophy, it is not pervasive and longstanding across the board. Even Kornblith, who defends a fairly skeptical account about progress in philosophy, appeals to recent consensus in the epistemology of peer disagreement in order to prevent self-defeat.²⁰ Similarly both Williamson (2006) and Stoljar (2017) produce a number of examples where philoso-

¹⁹ See footnote 8 on page 72 for a similar discussion.

²⁰ I am unconvinced that works of philosophy should avoid self-defeat, especially works such as Kornblith's that raise large-scale skeptical challenges for philosophy. I suspect critiques of the foundations of philosophy are necessarily philosophical in nature. This should not stop people from expressing those skeptical challenges on the grounds of the challenge's self-defeat. Otherwise we stifle critical analysis of philosophy's foundation. Instead, we should view non-consistent critiques of philosophy as worrisome *reductio ad absurdums* against philosophy's foundations. Nonetheless, in this chapter and the next I argue in a way that is self-consistent.

phers clearly know more than they did decades or centuries ago. While a reader may take issue with specific examples, some progress has certainly been made in philosophy.²¹ Such progress has many forms. Contemporary philosophers know more than historical philosophers in the form of conditional claims about what sort of commitments views have to take on board, refinements of said views, advances that have long been taken up by other academic disciplines, new thought experiments, dissolution of old problems, and new tools and concepts to help with old philosophical problems (Chalmers, 2015).

Producer-focused metaphilosophers could then defend the role of testimony in these cases where there has been obvious progress. Nevertheless, while some progress has been made and peer disagreement should not press us towards full skepticism about philosophy, some care is needed to separate out when progress allows testimony to produce testimonial knowledge and when it does not. This is the task of the next chapter, and the answer strongly points towards the importance of taking a *consumer*-focused approach to metaphilosophy.

3.6 Conclusion

In this chapter, I introduced producer-focused metaphilosophy, a metametaphilosophical commitment underlying exegesis-based intuition denial. The experimental attack, at least in its extant form, uses data about the consumption of thought experiments in order to draw skeptical conclusions about philosophy. In contrast, text-based exegesis attempts to block these skeptical conclusions by drawing inferences from data about philosophy's production.

I argued that philosophical testimony is the most promising defense of producer-focused metaphilosophy. This route faces two adequacy challenges. It needs to give an

²¹ Both authors are a bit too eager to claim progress in certain cases. For example, Stoljar (2017, 57) cites a passage in Goodman (1983) as his main evidence that Goodman (1983) dissolved Hume's problem of induction.

adequate epistemic story of how philosophers learn from each other and an adequate descriptive story that captures the actual practice of philosophy. In the final section of this chapter, I argued that peer disagreement raises a serious barrier to answering the epistemic challenge.

In the next chapter I turn to the descriptive adequacy challenge. I argue that actual practice tells against the widespread use of testimony in philosophy. Consequently, actual practice weighs heavily in favor of the metaphilosophical importance of the study of the consumption of philosophy. Moreover, actual practice weighs heavily against the metaphilosophical relevance of exegesis.

Chapter 4

How Philosophers Learn From Each Other

As I stated at the very beginning of Chapter 1, this dissertation has two goals. One goal is to critique exegesis-based intuition denial, and the other goal is to defend the use of a wide-range of empirical data in the study of the epistemology of philosophy. The last two chapters have focused primarily on the former goal. Chapter 2 argued exegesis-based intuition denial does not effectively defuse the experimental attack. Chapter 3 looked at how the method of exegesis could be defended.

Along the way, I began to formulate my answer to what empirical data matters to the epistemology of philosophy. Chapter 2 relied on empirically-testable claims about the history of philosophy, and Chapter 3 started an argument against the relevance of exegetical data. This chapter finishes the criticism of exegesis by showing that it has no place in the epistemology of philosophy but instead highlights what data we should pay attention to. Understanding how philosophers learn from each other highlights the metaphilosophical importance of data from anthropology, sociology, and new directions in psychology.

4.1 Returning to Producers and Consumers

In the last chapter I introduced the distinction between producer-focused and consumer-focused metaphilosophy. Exegesis-based intuition denial is a producer-focused metaphilosophy, since it prioritizes data about how thought experiments were created. Experimental philosophy is in contrast a consumer-focused metaphilosophy since it argues on the basis of how people react to thought experiments. In the last chapter I criticized existing defenses of producer-focused metaphilosophy (Section 3.3) and offered up testimony in those arguments' stead (Section 3.4).

As discussed in Section 3.4, testimony would provide a plausible defense of producer-focused metaphilosophy of the sort recently defended or assumed by metaphilosophers. If consumers of philosophy gain knowledge (or justification/warrant) through testimony, then learning about the methods and epistemic states of the producer will tell us something – but not everything – about consumers' epistemic standing. In response to this, I raised two challenges for producer-focused metaphilosophy. The *EPISTEMIC ADEQUACY CHALLENGE* requires producer-focused metaphilosophers to defend their position in a way that gives philosophers some way of learning from each other. I argued in Section 3.5 that peer disagreement makes meeting the epistemic adequacy challenge very difficult since most potential testimonial justification, warrant, and knowledge in philosophy is defeated.¹

The problem for producer-focused metaphilosophy's ability to meet the epistemic adequacy challenge is that when works of philosophy are produced, they are born into peer disagreement. That papers defend contentious philosophical views is, if anything, a virtue of those papers. Philosophical works are often judged based on whether they defend a novel position or whether they defend a contentious position in a novel way, and often part of what it is to produce good work in philosophy is to competently defend

¹ Much as in previous chapters, I hereafter drop talk of justification and warrant in favor of knowledge. Readers may replace talk of knowledge with talk of justification or warrant if they prefer.

something peers disagree with. Accordingly, a new work of philosophy is exactly the sort of work that cannot produce testimonial knowledge. Even if a work of philosophy explores something so novel no one actually disagrees with it, because of philosophy's history of widespread but not universal peer disagreement, we have reason to believe that relevant experts will soon disagree with the work. We cannot gain testimonial knowledge from novel and controversial works of philosophy, and this is a problem for producer-focused metaphilosophy's ability to meet the epistemic adequacy challenge.

The focus of this chapter is to push the other adequacy challenge, the DESCRIPTIVE ADEQUACY CHALLENGE, by developing an understanding of how philosophers gain knowledge from works of philosophy. There has to be some way we do. If testimony is the only way philosophers can gain knowledge from the works of other philosophers, there would be little epistemic gain in leaving the comfort of our own thoughts to engage with any other philosopher. Since we reasonably believe (or know that as a matter of fact) other philosophers disagree with an author's conclusions, it would be arbitrary to believe that author over others.

As discussed in relationship to the Moorean toy argument, we have a non-testimonial pathway to come to know a communicated proposition (or its negation) in philosophy. We can check an author's arguments and assertions by our own lights. Here I introduce this epistemic pathway and discuss its consequences. Section 4.2 introduces what I call SHOWING, where a work causes a belief but is not the epistemic basis of the belief. Section 4.3 argues that the prevalence of showing in philosophy is bad news for producer-focused metaphilosophy. Section 4.4 illustrates my account by applying it to the phenomena of hoax philosophy, specifically the 2018 "Grievance Studies" hoax. I argue that since science relies on testimony in a way philosophy does not, insincerely written philosophy can still be epistemically valuable philosophy. Section 4.5 considers and rejects the objection that I am mistaken in thinking showing is a distinct pathway from testimony. Finally, Section

4.6 wraps up discussion of exegesis-based intuition denial by drawing lessons from the last few chapters about what sources of empirical evidence are necessary to tell a complete story of the epistemology of philosophy.

4.2 Showing

Most works of philosophy cannot produce testimonial knowledge because of peer disagreement, but we still learn from other philosophers. This is because, instead of testimony, works of philosophy enable knowledge in consumers by drawing attention to relevant reasons or evidence that the consumers have access to. In this way, works of philosophy are akin to pointing. Just as pointing to something can cause someone to know something by showing them (see Grice (1957)), works of philosophy enable philosophers to see things they might not otherwise have appreciated or noticed.² Philosophy is hard, and sometimes other people catch things we do not. We read their works in order for them to point out what we might otherwise miss, and we learn from each other without relying on testimony.

When our attention is successfully drawn to something through an act of pointing or showing, the producer's act causes our belief that P without providing our justification that P. When a friend points to an embarrassing stain on my shirt, I come to know I have an embarrassing stain because *I see it*. It is only in a causal sense, not epistemic sense, of "because" that I know about the stain because of her gesture. In other words, she enables but does not justify my belief about the stain. I am able to know about the stain on my shirt because the evidence needed for knowledge is within what we might call my EPISTEMIC ENVIRONMENT. Before my friend shows me the stain, I might not actually possess the relevant evidence to know I have a stain on my shirt, but if I were to look for such evidence I could find it.

² Hills (2020) calls this process "propagation" instead of showing.

A communication act, whether speaking, pointing, gesturing, or writing, can produce knowledge via showing whenever it draws someone's attention to evidence or reasons within the consumer's epistemic environment. Showing can happen whether the evidence in the consumer's epistemic environment involves a coffee-stained shirt or abstracta. Turning to thought experiments, a thought experiment may draw my attention to counterexamples that I did not know exist, creating knowledge via showing. After reading a thought experiment, I know a theory has a counterexample because (in a causal sense) someone presented the thought experiment to me, but I know the theory has a counterexample because (in the epistemic sense) I had an intuition or considered some argument. Like the stain before I looked at my shirt, the counterexample is in my epistemic environment.³ As a matter of fact, I might never have had the occurrent intuition or judgement that a counterexample to a theory exists, but if I had considered the counterexample, I would have known that the theory is false.

Armchair philosophy by and large allows for an act of communication to be treated as showing. To see this, consider the Moorean argument from the previous chapter. Here it is again:

I know that I have a hand. My hand is external to my mind, so if I know I have a hand, I know there is an object external to myself. Given closure of known entailments, if I know there is an object external to myself, I know there is an external world. Therefore I know that there is an external world.

Each part this argument is what we might call CONSUMER-EVALUABLE. I have written this (imperfect) Moorean argument out in several discrete steps and you, as my reader, are able to evaluate the truth of each step and the overall validity of the argument. You are able to ask yourself whether you do know that you have a hand. You can think through whether I am properly using closure (or "closure"). You can evaluate for yourself

³ For a related view, see Ichikawa (2016); Ichikawa and Jarvis (2013).

whether I have inadvertently defended a strong modus tollens against knowledge of having a hand. Because the information necessary to make these judgements is in your epistemic environment, the Moorean argument can be treated as showing rather than telling.

Not only *can* you as the reader treat the Moorean argument as showing, but to know anything from the argument, you *must* treat it as showing. Because there is peer disagreement among anti-skeptics about whether Moorean solutions work, anything you learn from the toy argument is not from testimony.⁴ Instead, if you know anything because of the argument (in the epistemic sense of “because”), it is because you are mustering your own capacity to think philosophically – and that knowledge is based on your own competence as a philosopher, whatever that competency consists in.⁵

The prevalence of showing does not mean that philosophy is solipsistic. While our justification is our own, we work as community to enable knowledge in each other. The Moorean toy argument illustrates this communal enabling well. My ability to put together an anti-skeptical argument with this sophistication is the byproduct of an intergenerational project of knowledge production (see Bealer, 1998). Consider the refutation of Berkeley’s idealism when Samuel Johnson struck a stone and said “I refute it thus”. Johnson’s 19th century argument is very much in the same spirit of the Moorean toy argument, in that it argues for the external world by appealing to the obviousness of external objects. Moore built upon this by making the argument an explicit syllogism. Since Moore, epistemologists have developed Moore’s argument by, among other things, clarifying the role of closure in the argument.⁶

⁴ Arguably I am mentioning rather than using the Moorean argument. This adds another barrier to testimonial knowledge here.

⁵ The reader might worry there is tension between my optimism towards showing and my claim that peer disagreement defeats testimony. If the reader is not willing to take on board a view common in the peer disagreement literature that we should weigh our own judgements above those of others (see especially Huemer, 2011; van Inwagen, 1996), then I invite the reader to consider whether showing is the plausible primary source of interpersonal knowledge exchange in philosophy independent of whether testimonial philosophical knowledge is defeated.

⁶ I do not take this to be a definitive history of Moorean anti-skeptical arguments. Rather it is one plausible story of how one strain of anti-skeptical thinking has developed.

This progress was not the product of testimony. Instead, philosophers worked to epistemically clear up the environment for other philosophers (see Goldberg, 2013). By doing their own investigations into skepticism, they were able to progress the debate by *showing* other philosophers what the important aspects of skepticism are. This allows philosophers to learn from each other, even when testimony is not involved. However, this does mean that our justification *is* solipsistic. While we participate in a community of knowledge producers that communicate as a means to enable each other to form true beliefs, our justification is ultimately our own. Our beliefs epistemically stand or fall based on what justification we are able to muster for our own beliefs.

Notice however that testimony and showing are mutually compatible and potentially interlocking sources of knowledge. When we evaluate texts by our own lights, we may ultimately discover additional defeaters for potential testimonial knowledge. Grappling with a text often reveals reasons that the author does not seem to take into consideration, including false assumptions and bad reasoning. When we find these, we have source-sensitive defeaters (Casullo, 2003; Constantin & Grundmann, 2018) that remove reasons to believe the conclusions of the paper on the basis of the authors' testimony. This source-sensitive defeat does not defeat the evidence itself, but it defeats the connection between the evidence (*that the text says such and such*) and the proposition (the conclusion defended in the paper). Discovering that an author made a mistake does not defeat our own non-testimonial evidence for believing a proposition nor does it defeat our evidence that the philosopher said what they said. Rather, it defeats our justification for the proposition from our evidence that *the author said such and such*.

4.3 Exegesis Reveals Next to Nothing

The prevalence and necessity of showing raises a problem for producer-focused metaphilosophy: it is unable to meet the descriptive adequacy challenge raised in the previous

chapter. This problem for producer-focused metaphilosophy has two forms. First, relying on testimony to justify the method of examining original texts is a problem if such texts rarely produce testimonial knowledge. This first version of the descriptive adequacy challenge is specific to my discussion here and in Chapter 3 since it only bites if producer-focused metaphilosophy is defended on the grounds of testimony. The second form of the descriptive adequacy challenge is neutral on the point: the primacy of showing in philosophy paints a picture of the social epistemology of philosophy in which exegesis-based arguments about the epistemology of philosophy have no motivation. The *actual* epistemic states of the producer are largely epiphenomenal to the epistemic states of everyone else. Instead, as I discuss in this section, the social epistemology of philosophy reveals we should put our focus squarely on the consumer.

When consumers check for themselves, it is their own methods that matter to the justification of their beliefs. When an act of communication acts as showing instead of telling (i.e., testifying), the actual justification and knowledge of the author is irrelevant to the consumer's epistemic standing, especially if any putative testimony warrant or justification is defeated. Rather, the justification that the consumer ends up with will depend, if only partially, on the methods the consumer *thinks* the producer was using. For example, if an author intends to defend P on grounds G, and the consumer misreads the author as defending P on grounds G', and the consumer evaluates the argument on grounds G', then the consumer can still know P if grounds G' are good grounds.

Remember that exegesis-based intuition deniers are committed to the claim that philosophers have been widely mistaken about what the methods of philosophy are (Section 2.4.1). This is because denying intuition's role in philosophy requires prioritizing exegetical data over other philosophers' understanding of the methods of that text. Therefore, such intuition deniers still have to say that philosophers *think* original texts rely on intuitions. Combining that claim with the claim that philosophers learn from each other

via showing rather than telling means there is little to no reason to examine original texts of thought experiments. If consumers think that thought experiments' verdicts are justified by intuitions and in fact use intuitions as justification for their belief, then their justification relies on their intuitions. It does not matter if Gettier uses arguments to justify the verdict if everyone else uses intuitions. Everyone else's knowledge is based on intuitions, not arguments.

4.3.1 Consensus in Philosophy

This argument against producer-focused metaphilosophy only works if those original texts do not produce testimonial knowledge in readers, and my argument against testimonial knowledge relies on the prevalence of peer disagreement in philosophy. Nonetheless, it is not obvious that Gettier's paper, for example, is subject to peer disagreement, especially now. Perhaps it was reasonable in 1963 to expect peers to disagree with the paper's thesis, but it is not obvious readers should think that now. Since few if any epistemologists working in the 21st century disagree with Gettier, perhaps then some philosophers do know JTB is false because of testimony. This observation does not save producer-focused metaphilosophy. While some people likely have testimonial knowledge that the JTB account is false, we need to pay careful attention to where such testimonial knowledge comes from. It is not from Gettier.

Sometimes philosophical experts do not disagree and instead come to a consensus about a work of philosophy. Gettier is widely acknowledged as having a successful counterargument to JTB, and Foot's trolley case and Thomson's later variations are widely considered to successfully pick out a crucial fault-line between deontology and consequentialism. Lewis's *On The Plurality of Worlds*, while not widely considered sound, is considered by expert commentators to be extremely well-argued. Descartes' argument for knowledge of the external world is widely accepted to go wrong by the time he defends

the existence of God. There will always be the odd expert who disagrees with these received views, but they are nonetheless received views among the vast majority of relevant experts.⁷ Because of this consensus, we lack defeaters from peer disagreement and can know these propositions through testimony.

When we take such claims on testimony, our knowledge is based on the testimony of those who themselves considered the text by their own lights. In instances where works are accepted as good, defeaters from peer disagreement go away when experts form a warranted consensus about a work. The experts could not have known from the producer through testimony because they had defeaters for testimonial knowledge while consuming the text. The author merely played an enabling role for the consuming expert, who then testified the work's goodness to others. The justification for successful philosophical testimonial knowledge does not depend on the actual intention or methods of the *producer* of the work. It depends on what experts considered themselves, which depends in part on what the experts *perceived* to be the intention or method of the author and what other considerations they independently explored. Therefore, even when philosophical testimonial knowledge about a work's thesis is possible, at no point in the story does the actual methods or epistemic standing of the producer play a role "downstream" from the work.

Because testimony is coming from experts and not the original producer, misinterpretations (such as thinking there are intuitions when there are in fact arguments) are not in themselves problematic. When consumers evaluate the arguments of original texts, consumers can fail to understand the true nature of a original work of philosophy or the producer could make a mistake while communicating their ideas. A producers may misstate a step of an argument, leave a premise out, or base a claim on an idiosyncratic and unstated assumption. Sometimes this will be unlucky where, despite the producer's

⁷ I am not going to try to adjudicate what proportion of peers need to agree to enable testimonial knowledge. For present purposes, I am assuming it is somewhere between 50% and 100%.

knowledge, consumers widely reject the argument. Mistakes could also be lucky, leading consumers to learn something from consuming the text that the producer themselves did not know or understand. Testimony based on this mistake would transmit the mistake, not the actual methods of the work. Therefore, if experts think Gettier relied on intuitions, any testimony of Gettier involves justification based on intuitions.

The exegetical projects of Deutsch and Cappelen on their own do not tell us anything about the epistemology of anyone except the original authors. The only philosopher whose knowledge is based on the epistemic states of the producer is the producer themselves. The philosophical knowledge we have based on texts depends either on our own evaluation of the perceived argument or undefeated testimony of someone who has. How then do we know how a philosopher stands epistemically in relation to a proposition defended in a work of philosophy? If they have testimonial defeaters, then we need to look at how *they themselves evaluated the argument*. If on the other hand they lack defeaters for testimony, then we need to look at how the testifiers evaluated the arguments for themselves. At each point, metaphilosophical epistemic evaluation requires that we look at consumers of the text, not the producers.

4.4 Fake Philosophy is Real Philosophy

The above picture of the social epistemology of philosophy has consequences beyond the debate between experimental philosophers and Deutsch and Cappelen. In this section, I further develop the picture of the epistemology of philosophy through a discussion of insincerely written philosophy.⁸ This also serves the purpose of further illustrating the distinction between showing and telling.

Over the course of 2017 and 2018, seven hoax humanities papers written by a team of two writers and one professional philosopher were accepted by peer-reviewed journals

⁸ Thanks to Joseph Bowen for suggesting the connection between my work and this topic.

(Pluckrose, Lindsay, & Baghossian, 2018). Targeting contemporary structuralist thought, specifically feminist theory, the team wrote papers that sounded like contemporary works of feminist theory but that defended conclusions the team thought were absurd. The 2018 “Grievance Studies” hoax, named after the hoaxers’ pejorative name for the disciplines they targeted, gained some notoriety in right-leaning publications and social media as ammunition against academia. Fox News posted an article online titled “Academic journal accepts feminist ‘Mein Kampf’” and said the authors “point out this sort of ‘biased research’ would not be accepted in any other industry, but is pervasive in higher education” (Parke, 2018). Other conservative outlets drew parallels with #MeToo, saying, for example, that by “propping up absurd claims with the paraphernalia of respected scholarship – peer review, respectable journals, technical language and so on—[scholarship in the target areas] provides the Twittermob with the pitchforks it needs to lynch its targets” (Cook, 2018).

Scholarship on past academic hoaxes has focused on the role of so-called gatekeepers in academia.⁹ Editors, reviewers, and ethics committees have a disproportionate say in what research becomes available for broader consumption by other academics and lay people. While early commenters on the Grievance Studies hoax have focused on such gatekeepers and what the hoax actually demonstrates about the targeted disciplines,¹⁰ little has been said about the epistemic consequences of the publication of hoax papers. This section explores what happens when people read papers that are meant to discredit or damage the field in which they are published. As it turns out, in certain limited cases, including the recent Grievance Studies hoax philosophy article accepted by *Hypatia*, insincerely written scholarship can further progress in a field, even if readers know it is insincerely written. Because philosophical scholarship typically shows the reader its claims as opposed to testifying them, hoax scholarship can produce knowledge of a paper’s thesis in its readers

⁹ For scholarship on past hoaxes, see Baghossian (1996); Goldman (1999); Kuroki (2000); Spera and Peña-Guzmán (2019).

¹⁰ E.g., Hanlon (2019); Phipps (2018); Soar (2018).

– even if the author(s) do not believe the thesis.

4.4.1 Hoax Science

To draw out why hoax philosophy can still produce knowledge, it is helpful to contrast how philosophy papers transfer knowledge to how scientific papers transfer knowledge.

Scientific papers typically have five distinct sections with five distinct roles in the paper. The INTRODUCTION introduces relevant literature as a way of providing necessary background for the paper. The introduction is also often used to justify the importance of the questions the experiment is examining. Then, the MATERIALS AND METHODS (or just METHODS) section explains the experiment's procedures. Following this, the RESULTS section presents the findings of the experiments explained in the methods section. Rather than presenting raw data (which is increasingly available online), the results section presents the results of statistical analyses. The DISCUSSION section then draws inferences from the data and places the findings in the larger literature discussed in the introduction. The CONCLUSION ends by briefly summarizing the paper.

The reader's epistemic relationships to the methods and results sections are very different than the reader's epistemic relationship with the discussion section. In both the methods and results section, the reader has no choice but to trust the authors that what is written reflects what actually happened. Historically, the vast majority of hoax science has occurred at this stage. Methods, materials, or findings are either created wholesale or subtly manipulated in order to support a given conclusion.¹¹ Hoaxes are common here because the claims in the methods and results sections are outside of the reader's epistemic environment; it is beyond the reader's ability to check the fact of the matter for themselves. Consequently, insofar as the reader knows the experimental results, they know through

¹¹ Note that I am not including p-hacking or other statistical manipulation here. Despite appearing in the results section, the epistemic features of statistical manipulation are the same as claims in the discussion section, which I discuss below.

testimony.¹² Note, however, that this is not to say that the reader cannot spot defeaters for the contents of these sections, as wildly infeasible methods or improbable findings can be reasons not to believe what is testified.

Compare this to discussion sections, where inferences are drawn from claims made earlier in the paper. The claims in the discussion section can be understood as conditional claims: given previous research and the research original to this paper, P. Unlike the data presented in results section, readers are well within their epistemic power to evaluate the inductive or abductive validity of such inferences. Unlike the methods section, where readers have to take the word of the authors, readers have access to sufficient evidence to evaluate the claims being made in the discussion. They can check for themselves if the conclusions drawn in the discussion section are actually supported by claims made previously in the paper.

Superficially, the methods, results, and discussion sections of a scientific paper resemble each other. In all cases the authors are asserting claims that they are experts about. Presumably no one knows the potential inferences to be drawn from authors' experiments better than the authors. Nonetheless, the relationship between author and reader is very different depending on where in the scientific paper the assertion is made. In the material and methods and results section, the reader's knowledge of the asserted proposition must be based on the testimony of the author. Even if the reader replicates the experiment, the propositions expressed in the materials and methods and results section, that the experimenter did x and found y, is only known through testimony. Later, in the discussion section, the testimony is impure.

¹² As more and more raw data is posted online, the amount that is taken on trust is shrinking. For some papers it may be possible to upload videos of data collection, for example. These instances are rare, and even then it is taken on trust that the videos show the whole story and are not hiding manipulation on the part of the researchers.

4.4.2 Hoax Philosophy

If the epistemic dynamic between reader and writer of a work of philosophy is similar to the relationship between reader and writer of the methods and results sections of scientific papers, then the epistemology of the Grievance Studies hoax philosophy paper accepted by *Hypatia* is fairly straightforward. As readers, our knowledge of data presented in scientific papers is testimonial. Insofar as we know the facts of an experiment – its methods and results – our knowledge is based on the communicative act of the author writing the paper.¹³ Consequently, as a reader, our knowledge of the experiment is fragile and easily defeated.

Earlier in this chapter, I argued while knowledge can be *caused* by engaging with other philosophers, it is not epistemically *based* on the work or epistemic standing of those philosophers. To illustrate why the difference matters here, consider FELINE GETTIER, which is stripped of the distracting political elements of the Grievance Studies hoax:

In the early 1960s, a man decides to write a prank academic paper. The man chooses words from then-contemporary epistemology papers, including “justified”, “the”, “knowledge”, “know”, and “Jones”. He writes these words on pieces of paper, sets the pieces of paper on the floor, and puts pieces of fish on each paper. He then sets his cat in the middle of the floor, letting the cat wander from piece of fish to piece of fish. As the cat eats a piece of fish, the man writes down the corresponding words in order on a piece of paper. Once the cat has its fill of fish, to the man’s delight, the resulting string of words looks uncannily like actual epistemology. With this in mind and keeping the words untouched, the man formats the result, adds punctuation, and with a snicker, submits the resulting paper to *Analysis* under the name of his cat, “Edmund Gettier”.

¹³ See Lackey (2008) for the distinction between testimony being based on the speaker’s epistemic states compared to the communicative act.

The editor at *Analysis* sends the paper off to a referee and receives a report back stating that the argument is pithy, succinct, and apparently sound. The referee even reports that the paper has convinced her that JTB is not sufficient for knowledge, despite her previous beliefs to the contrary.

The man who submitted the paper to *Analysis* comes clean and admits what he did. The editor passes this on to the referee, who responds to the editor by saying “So what? I now know that JTB is not sufficient for knowledge.”

Despite having potential testimonial knowledge defeated by the author’s insincerity, the referee nonetheless maintains her knowledge of the proposition defended in the paper upon learning of the hoax. To see why, consider the argumentative structure of Gettier (1963).¹⁴ First, a view – the JTB account of knowledge – is introduced. Then, two counterexamples are presented. Finally, the paper infers from the counterexamples that JTB is not sufficient for knowledge. Unlike someone else’s scientific data, which we need to depend on testimony to know, we do not need testimony to know the various steps of the Gettier argument. The two thought experiments elicit intuitions *in the referee*, so the referee has independent access to the justification of the corresponding step in the argument. Similarly, whether the verdicts of the case are actual counterexamples to the sufficiency of the JTB account of knowledge is well within the referee’s power to judge for herself. Even if, following Deutsch (2015), Cappelen (2012), Williamson (2007), and Ichikawa (2016), intuitions do not play a role here, whether the character in Gettier cases knows is in the epistemic common ground of the author and reader (Weinberg (2014); Machery (2017, 178)), so the referee still has it within her power to judge for herself.

The ability of the reader to evaluate the claims being made in the text independently of that text is not unique to philosophy. It is available to the reader whenever they have independent access to propositional justification for what is being asserted in the text.

¹⁴ Please note that I am explaining why the referee knows, not arguing that the referee knows.

This means that even in distinctively empirical disciplines knowledge can be gained from reading hoax papers – if the hoax only affects certain parts of the text. In scientific papers, while the reader does not have access sufficient propositional justification to autonomously evaluate experimental data, they can evaluate conditional inferences made from experimental data (assuming they know the data from some other source). So while hoax scientific data is epistemically troublesome, hoax scientific *arguments* need not be.

The “Grievance Studies” hoax paper in *Hypatia* contains no original empirical work. It does, however, insincerely argue that hoax papers targeting feminist critical theory (e.g., the very paper making the argument) are immoral. Whether this thesis is true, readers do not need to rely on the authors’ testimony to know this. Indeed, like every other novel work of philosophy, readers can learn from the hoax text because the text enables readers to consider philosophical arguments and reasons they might not have thought of on their own. It does not matter if the text is written insincerely because readers can independently use their own epistemic tools to consider the arguments in the text. Therefore, readers may find the hoax work enlightening (as I did) even knowing the work was insincere.

While *Hypatia* retracted the hoax paper, they lacked an important and merely apparent reason to do so. Unlike the fraudulent scientific data in other Grievance Studies papers, justification readers have for their philosophical arguments is not defeated by the hoax. So while perhaps the hoax paper is embarrassing for the journal, the continued availability of the hoax article does not preclude readers from learning things from the paper’s arguments. Instead, the paper would have continued to enable readers to use their own epistemic tools and explore their epistemic environment in order to adopt or reject the claims made in the text. Regardless of the intentions of the authors, the fake philosophy may still have something to offer to the rest of us.

4.5 Objection: “Showing” is Just Testimony

In this and the previous chapter, I claimed showing and testimony are two distinct epistemic pathways in which we learn from each other. In this section I consider the objection that all the cases I consider of knowledge caused by works of philosophy are actually just testimony. Therefore, according to this objection, there is a strong epistemic link between producers and consumers and hoax philosophical arguments are as problematic as hoax scientific data.

This objection cannot explain why the referee in Feline Gettier knows JTB is false. Most accounts of what makes a communicative act testimony place epistemic requirements on the creator or creation of the act. On these accounts, the paper in Feline Gettier fails to be genuine testimony. Other accounts do not place such requirements on the communicative act, but on accounts of testimony where Feline Gettier is genuine testimony, reading the paper fails to produce testimonial knowledge, justification, or warrant.

To respond to this objection, I look at three notable accounts: Coady (1992), Sosa (1991), and Lackey (2008). Both Coady and Sosa’s accounts require certain conditions on the producer for something to be testimony in the first place. Therefore, Coady’s and Sosa’s views entail that the referee lacks testimonial knowledge by denying that testimony occurs. In contrast, Lackey defends a much more open and permissive analysis of testimony, and on her account, testimony – at least in one sense – occurs. Nonetheless for testimony to be justified, Lackey places substantive requirements on production, and so on her account the referee’s belief is not justified by the text even before the referee learns about the cat. Thus on all three accounts, the referee lacks testimonial knowledge.

Sosa’s account of testimony is the simplest of the three considered here, stipulating that someone testifies that p iff they state their belief that p (1991, 219). It is unclear who is making the statement in Feline Gettier. In a sense the cat is making the statement because it inadvertently chose the words. In another sense the man is making the state-

ment because he wrote the argument down and sent it to a journal. Regardless of whether we understand the human or the cat as making the statement, no testimony occurred on Sosa’s account. Neither the cat nor the man believe the contents of the paper sent off to *Analysis*, so neither testified that justified true belief is not knowledge.

Coady’s (1992) analysis of testimony is slightly more complicated than Sosa’s, placing three jointly necessary and sufficient conditions on something being testimony. Nonetheless, for purposes here we can focus on one particular necessary condition, namely, that for an act to be testimony that *p* by some speaker *S*, it must be the case that “*S* has the relevant competence, authority, or credentials to state truly that *p*” (Coady, 1992, 42). This condition places a requirement on the creator rather than on the consumer’s perception of the creator. To be an act of testimony on Coady’s account, the producer of the act of communication must be in some privileged epistemic position relevant to *p*. The producer of the act of communication in Feline Gettier, whether understood to be the man or the cat, clearly lacks this privileged epistemic position. Therefore the act of communication fails to be testimony on Coady’s reading.

Because Sosa and Coady place epistemic requirements on a speaker in order for the speaker to genuinely testify, the communication act in Feline Gettier fails to be testimony on their accounts. Lackey (2008) takes issue with such accounts, arguing that the question of whether testimony occurs is ultimately a metaphysical question, not an epistemic question. By placing epistemic requirements in the analysis of testimony, Lackey argues, authors like Coady and Sosa are mixing up issues of epistemology with issues of metaphysics. Accordingly, Lackey argues that testimony can occur independently of the speaker’s epistemic states.

Lackey proposes a disjunctive account of testimony, separating testimony into two distinct types – testimony from the point of view of the speaker and testimony from the point of view of the hearer. Talk of a single type of testimony, Lackey argues, obscures

the dual nature of testimony. There is a sense in which a speaker testifies that *p* even if no one is paying attention or believes them, and there is a sense in which a speaker testifies that *p* even if the comment was meant to be private but is overheard by someone.

To capture both notions, Lackey defends a disjunctive account of testimony:

S testifies that *P* by making an act of communication *A* if and only if (in part) in virtue of *A*'s communicable content, (1) *S* reasonably intends to convey the information that *P* or (2) *A* is reasonably taken as conveying the information that *P* (2008, 35-36)

While we may often think of testimony from the producer's prospective as Coady and Sosa do, Lackey argues that producer intention, expertise, or belief is not necessary for an act to be testimony. Private documents found posthumously, overheard conversations, and predictable lies can all serve as testimony. If I know someone is going to lie to me about the answer to a yes/no question, I can use their lie as testimonial justification by taking the opposite of their answer as being true.

Each of the two disjuncts is meant to capture a distinct notion of testimony – SPEAKER TESTIMONY and HEARER TESTIMONY. Lackey's speaker testimony takes the point of view of the producer and will not be relevant here since the speaker in this case does not take themselves to be communicating information. Hearer testimony, on the other hand, purposely does not make any reference to the conditions giving rise to the speech act, instead focusing on how it is perceived:

Hearer Testimony: *S* *H*-testifies that *P* by making an act of communication *A* if and only if *H*, *S*'s hearer, reasonably takes *A* as conveying the information that *P* (in part) in virtue of *A*'s communicable content. (Lackey, 2008, 32)

On Lackey's view, the act of communication itself carries the epistemic value, which is why on hearer-testimony the hearer reasonably takes the act itself as conveying information. This language is meant to be fairly broad to capture non-verbal acts (such

as nodding) as well as pragmatic implicature (such as answering “is he a good cook?” with “he sure tries” – the implicature being that he is not). While Lackey keeps the applicable communication acts broad, she does want to exclude certain acts that were not intended to express communicable content. These include instances where a non-communicative act is mistaken for one or apparent communicative acts turn out to be coincidental and the byproduct of random chance. In this way, Lackey’s hearer-testimony does place a restriction on the origin of the act, namely that the producer created an act of communication. (Lackey, 2008, Ch1 fn31)

Despite hearer-testimony’s focus on the consumer’s takeaway from the act of communication, there is some difficulty here focused around whether an act of communication actually occurs in *Feline Gettier*. To be an act of communication, Lackey requires that the producer of the text intends to express communicable content. Instead of providing analysis or stipulating a definition of what it takes to intend to express communicable content, Lackey communicates the notion with examples. She argues a private diary and private soliloquys intend to express communicable content, even if the producer of the communication intends no one to hear it. Thus, if someone consumes the private acts, either by overhearing the soliloquy or finding the journal, what they consume acts as hearer-testimony. Compulsive tics and random natural events interpreted as communication, such as Lackey’s example of branches on an island randomly spelling “HELP” to passing airplanes, do not, on Lackey’s view, count as testimony because they lack the intention to convey communicable content.

In some ways, the potential communication act in *Feline Gettier* reflects Lackey’s examples for both intentions to express communicable content and her examples of cases that do not. The cat did not intend to express communicable content when eating fish – in a way the situation is no different than randomly scattered branches spelling “HELP” – but the cat’s eating choices are not the entirety of the ostensible communication act.

The man chose which words the cat could have eaten, and the man also recognized and sent out the paper as an apparent work of philosophy. This last step is key to see that the prankster in Feline Gettier did intend to express communicable content. Even if the origin resembles “HELP” in random branches on the beach, the man recognized the text as containing communicable material and sent it out intending to express it.¹⁵

Granting that the transfer of information in Feline Gettier counts as testimony on Lackey’s view, we still do not have an account of the referee’s justification after finding out about the paper’s origin. Even given Lackey’s consumer focused account of hearer-testimony that feeds into her disjunctive account of testimony, Lackey’s account of testimonial justification still places requirements on the communication act’s origin. Because the paper in Feline Gettier was written insincerely, it fails to provide the referee with testimonial justification. Lackey provides a number of necessary conditions for testimonial justification, including the requirement that testimony must be “reliable or otherwise truth conducive” (Lackey, 2008, 176-177). Because this requirement refers to the act itself rather than the consumer’s understanding of it, on Lackey’s view of testimonial justification, the referee lacks justification from the start.¹⁶ The paper’s text was true only because of luck, not because the production had any connection to the truth of what knowledge is.

Therefore, on all three accounts considered here, the hoax philosophy paper does not produce testimonial knowledge despite the its ability to cause knowledge in the reader of the communicated propositions. While not comprehensive, the discussion in this section offers insight into why Feline Gettier does not involve successful testimonial knowledge on other accounts. If the metaphysics of testimony depends on the epistemic properties

¹⁵ This is a somewhat strange byproduct of Lackey’s view. Notice that this means that the paper is still testimony after the referee finds out about the cat. Because even if the referee learns that the paper was written insincerely, she can still reasonably take the prankster as conveying information that p in virtue of the communicable content of the paper.

¹⁶ Lackey has other necessary conditions that preclude justification in Feline Gettier, including that P has no undefeated defeaters for A’s testimony (2008, 177).

of the speaker or listener, insincerely produced philosophy will fail to be testimony. If, on the other hand, the metaphysics of testimony is neutral on an act's epistemic properties, requirements placed on testimonial knowledge will preclude the consumer of the hoax from having testimonial knowledge. Hoax philosophy or other insincerely philosophy does not produce testimonial knowledge, but philosophers can still gain knowledge of the theses the works appear to defend. Showing is distinct from telling.

4.6 The Importance of Empirical Facts

Now that I have completed my discussion of the social epistemology of philosophy, I can now tell a richer story about what empirical data is relevant to the epistemology of philosophy than I could at the end of Chapter 2. My argument in Chapter 2 turns on a single claim about the history of philosophy. My argument works only if the influential works of philosophy discussed by exegesis-based intuition deniers did in fact change other philosophers' beliefs. I argued Deutsch and Cappelen get in trouble because they do not consider how this historical influence interacts with their claim that their analysis of texts reveals nondecisive arguments. The failure to appreciate historical context is not trivial. The single quotidian fact that influential thought experiments are influential is enough to render an entire strand of arguments in the epistemology of philosophy unpalatable.

Paying attention to the history of philosophy while studying the epistemology of philosophy is hardly revolutionary. Deutsch himself compliments his exegesis of Gettier by looking at the language of early works of Gettierology (2015, 89-94). Brown (2017) criticizes Deutsch by arguing that denying intuitions' role in Gettier cases cannot explain why epistemologists in the 1960s thought Gettier was right despite disagreeing and changing their minds about why. Colaço and Machery (2017) point to historical evidence that epistemologists circa 1963 were interested in the concept of KNOWLEDGE (not knowledge itself), arguing intuitions would therefore have been playing a role in such conceptual

analysis. Indeed, metaphilosophically-motivated historical analysis of the use of thought experiments and Gettier's paper predate the debates I have been engaging with here (Hintikka, 1999; M. Kaplan, 1985).

While all of these analyses are on the right track, the present discussion of the social epistemology of philosophy allows us to say something more meaningful about what sort of historical analysis is important to the epistemology of philosophy. While, as I discussed in detail in Chapter 2, facts about *which* beliefs changed are important, it is also important to study individuals' chosen processes for changing those beliefs. In particular, we need to study the epistemic attitude philosophers take towards reading the work of other philosophers. Consumers of philosophy have the opportunity to approach a work of philosophy as showing *or* telling them the work's thesis. Depending on the circumstances one, the other, or both can produce knowledge.

This raises two distinct empirical questions. First, as a matter of historical and contemporary fact, what epistemic norms do philosophers have about approaching a work of philosophy? Second, in what circumstances do the epistemic tools that philosophers in fact use lead to epistemically good or bad outcomes? I now look at both questions in turn.

4.6.1 Philosophers' Norms

To understand which epistemic tools philosophers use when reading the works of others, we need to draw upon the methods of anthropology and sociology.

The last two chapters discussed the difference between showing and testimony. These are not mutually exclusive pathways to knowledge. Necessary and sufficient conditions may be in place for me to know based on testimony, but I may still check the truth of a philosophical claim by my own lights. Importantly, this choice is my own. Because most claims in philosophy papers are consumer-evaluable, consumers have the choice about

whether to evaluate the argument by themselves or take the claims on the word of the producer. It is an empirical matter which one philosophers choose, and it is not the sort of empirical matter that can be easily tested by existing experimental philosophy paradigms.

Instead, what attitude philosophers take towards a text is the sort of question that anthropology and sociology are well-suited for. The question of which attitude philosophers take towards a text is partially a issue of norms. It turns on how they were socialized as philosophers, specifically what norms they internalized about reading and listening to talks.

Analytic philosophers are generally a querulous group of people. Treating a text as testimony is probably fairly uncommon since that requires taking a producer at their word. Without further work, however, it is impossible to say how universally true this is. As I have discussed (Sections 3.4 and 4.3), it seems that contemporary analytic philosophers sometimes take claims on testimony, especially when the claim is beyond their area of specialization. Similarly, in some of the more obsequious corners of historical philosophy, philosophers likely believe historical figures' claims on testimony. Moreover, norms about this may have changed over the historical course of analytic philosophy, and any observations made about how analytic philosophers act now may not apply to analytic philosophers 50 years ago.

Sociology and anthropology give us the tools for checking for how philosophers approach texts. Sociology is the far more quantitative of the two and can provide a macro picture of how beliefs and norms affect analytic philosophy as a network.¹⁷ Anthropology is in contrast more qualitative and personal. Where sociology can tell us about the field as a whole, the anthropology of philosophy can reveal the implicit rules, expectations, and behavior driving members of our social groups. Focusing the methods of either discipline will therefore reveal something about the epistemology of reading and therefore help tell us the story of the epistemology of philosophy.

¹⁷ For an introductory – but qualitative – look at this, see Fuller (2002); Misak and Talisse (2019).

4.6.2 Appropriate Uses of Epistemic Tools

Sociology and anthropology will help reveal which epistemic tools philosophers use, but further empirical work is needed to reveal whether philosophers used the appropriate tools in the appropriate circumstances. They will tell us what philosophers did, but more empirical work is needed to tell us whether what they did produced knowledge.

Existing experimental philosophy tells us part of the story, since it reveals the epistemic value of treating *certain* works of text as showing. When we consider a thought experiment by our own lights, we have reactive mental states (e.g., intuitions). Experimental study of how people react to thought experiments then reveals whether those reactive mental states are good evidence for certain beliefs or not.

Thought experiments are a notable but small subset of philosophical communication, though. Lots of the works philosophers consider by their own lights are arguments, both formal and informal. A fuller experimental picture would therefore also analyze the way philosophers' beliefs change in the light of arguments and reasons. Unlike testing norms, this seems well within the reach of current experimental paradigms. For example, Wysocki (2017) found arguments following thought experiments did not change participants' responses about those thought experiments. More work in this direction will allow experimental philosophers to tell a richer story about whether philosophers' judgments about formal and informal arguments are subject to the same skepticism as philosophers' judgments about thought experiments.

Such experimental work nonetheless only tells us part of the epistemic story, since it tells us only about when philosophers treat a work as showing them something. When anthropologists and sociologists discover philosophical works are treated as testimony, we need to know whether the testimony has (undefeated) defeaters. This will partially depend on whether there was actual or reasonable expectation of disagreement among philosophers who are experts on the topic. Knowing this requires empirical study of the

extent of disagreement, both contemporary and historical. While some attempts have been made in this direction, especially the qualitative historical work of Stoljar (2017) and the quantitative contemporary work of Bourget and Chalmers (2014), more empirical study is needed before we can confidently say peer disagreement did not defeat putative philosophical testimonial knowledge in any one instance.

To summarize, because it is the epistemic standing of the consumer that matters to the epistemology of philosophy, metaphilosophers need to expand their empirical study of consumers. Sociology and anthropology will help tell us whether consumers treat a work of philosophy as testimony or showing. If it is treated as testimony, we need to study the extent of peer disagreement in order to build a fuller picture of what instances doing so produces testimonial knowledge of philosophical facts. If consumers treat the work as showing, then experimental methods from psychology will do. Not only will no one method work in all circumstances, but we need to combine multiple empirical methods from multiple disciplines to tell a complete picture of the epistemology of philosophy.

4.7 Conclusion

This wide-ranging chapter revolved around a central question. Given that peer disagreement defeats lots of philosophical testimony, what is the social epistemology of philosophy? My answer is that texts show consumers the truth or falsity of what is being argued. This was illustrated in contrast to scientific communication, where I used hoax philosophy to show that insincerely written philosophy can still produce knowledge in readers.

This discussion had two purposes. First, it finished a line of criticism of exegesis-based intuition denial started in Chapter 3. Epistemic states caused by notable works of philosophy are based on the epistemic processes of the consumer, not the producer. Therefore, exegesis of influential philosophy texts reveals little of epistemic note. Second, the positive story put forth in this chapter allowed me to defend a broad empirical program

in the epistemology of philosophy. In order to better understand the epistemic standing of philosophers, existing psychological paradigms should be broadened, our understanding of peer disagreement in philosophy should be expanded, and avenues in the anthropology and sociology of philosophy should be explored.

The rest of the dissertation will continue in the spirit of this latter point. Exegesis-based intuition denial will be largely set aside, although it will briefly reappear in Chapter 7. Instead, I focus on what other empirical data is needed for the epistemology in philosophy. Specifically, I defend the importance of the lexicography of philosophy in order to clarify existing debates.

Chapter 5

Polysemy

In the previous three chapters, I developed and explored a negative thesis. I argued that the exegesis-based arguments of Deutsch and Cappelen fall short, both on their own merits and methodologically. During these chapters, I began to argue for my positive thesis by showing how contingent facts about the field of philosophy are needed to answer questions about the epistemology of philosophy.

In the next two chapters, I set the negative thesis aside and focus on the positive thesis. In particular, I argue for the relevance and importance of lexicography to understanding the methods and epistemic status of philosophy. To argue this, I use “intuition” as a case-study. I put pressure on the idea that philosophers should be discussing intuition *simpliciter* because lexicographic evidence suggests that questions like *what are intuitions?* and *do intuitions justify beliefs?* do not have a single answer.

There is no single answer to these questions because “intuition” does not act like “water”, “tiger”, or “arthritis” are thought to, where a single meaning is discovered or determined by experts. Instead, the meaning of “intuition” is more like the history of “tweet”, “string”, or “square”. The verb “to tweet” was once an onomatopoeic term for imitating a bird, and now it is both that and a way the president of the United States reaches supporters. “String” once meant thin pieces of rope, but now it means that *and*

the fundamental entities posited by string theory. “Square” was once the name for the tool that carpenters and architects use to draw 90 degree angles, and now it is the name of that tool *and* the figure bound by 90 degree angles in Euclidean space. “Intuition” is no different. Over the course of its roughly 500 year history as an English word, philosophers have tried – and sometimes succeeded – at introducing new technical meanings of “intuition” in order to talk about mental state categories of theoretical interest. Rather than revising, replacing, or otherwise misusing “intuition”, the technical meanings of “intuition” that stuck now exist side-by-side with older meanings of “intuition”.

Defending this claim in Chapter 7 requires a fair bit of stage-setting. In this chapter and the next, I develop a metasemantic framework for technical language that highlights the importance of lexicography to the study of philosophy’s methods *and* enables the lexicography of philosophy as a practice. I argue that specialized discourse is full of ordinary words with non-ordinary meanings. This argument has two steps. In this chapter, I discuss POLYSEMY – the phenomena where words have multiple related and established uses – and argue that these represent distinct literal meanings of the word. In the next chapter, I discuss this view in relationship to technical language, or JARGON, and I argue that polysemy plays a common role in technical discourse. Many words have polysemous meanings that are only known by people with significant levels of technical knowledge. Once I have this in place, I move to Chapter 7 and argue that “intuition” looks to be following the path of such polysemous technical language.

5.1 Introducing Polysemy

While rarely discussed explicitly, there is a default assumption in philosophy of language, especially among semantic externalists, that the vast majority of words have one semantic meaning. While philosophers of language occasionally discuss homonyms, homophones, and homographs like “bank” (the financial institution) and “bank” (the side of a river),

these are treated as the exception rather than the rule. Call the received view that a word has a single meaning the *Univocal Assumption of Language*. While, to my knowledge, it is not explicitly defended anywhere, it pops up particularly clearly in recent discussion of semantic change.

Dorr and Hawthorne (2014) and later Cappelen (2018), have framed discussion of semantic change around “salad”, where “salad” allegedly originally meant mixes of cold vegetables but has slowly broadened in meaning over time to now include egg salad, tuna salad, and fruit salad. Both Dorr and Hawthorne (2014) and Cappelen (2018) frame this as a shift in a single meaning from a narrow sense to a broad sense. However, dictionaries that recognize the difference between the leafy-green use of “salad” and the broader use of “salad” (including Collins, the OED, and Merriam-Webster), typically label these two uses as contemporary, co-existing, and distinct definitions of “salad”.

In a similar vein, Sawyer (2018, 2020) has used “meat” as her preferred example of semantic differences over time. “Meat”, the thought runs, meant food in general when Shakespeare used it, but now “meat” only means animal flesh. Like “salad”, this history is also oversimplified. The OED3 – the world’s premier historical English-language dictionary – cites an examples of the “newer” flesh-sense of “meat” that predates Shakespeare by two centuries. The OED3 also cites Shakespeare as using the “newer” flesh-sense, the “older” food-sense, and three other senses of “meat”, including *the flesh of a fruit or nut*, and the figurative *source of strength*.

Philosophers’ discussion as if words have a single meaning that changes over time is in striking contrast with how other fields studying language, especially lexicographers and cognitive linguists, think about language. Some cognitive linguists take it as a given that words are associated with rich cognitive networks of interrelated metaphors and meaning (Lakoff & Johnson, 1980), and the overwhelming majority of dictionary entries for common English words contain more than one definition. “Bachelor” is not just defined

as a married man, it is also defined as both a person who holds a bachelors degree and the name of young male fur seals. Celestial bodies are stars, but so are 5-sided pointed figures, famous people, and the ratings of hotels and restaurants. As I write, I can see two separate types of things captured by different dictionary definitions for windows: a hole in the side of the building as well as the work area of my computer's word processor.

Unlike homonyms, which involve run-of-the-mill ambiguity between two semantic meanings of a word, polysemy occurs when words have multiple distinct patterns of use that are related.¹ This is extremely common with noun/verb pairs. Consider the following examples. The example of window/window, star/star, and bachelor/bachelor differ from standard homonyms like bank/bank in at least two ways. Not only do window-as-computer-workspace and window-as-hole-in-building share etymologies from when the window-as-computer sense arose in the 1960s, but the senses *feel* related. This felt relation has no doubt been strengthened by the allusion between the two in the name of Microsoft's Windows operating system and the OS's logo, which is a window. In contrast, "bank" (the building) and "bank" (the side of the river) were borrowed from separate languages at separate times and generally feel unconnected in the minds of speakers.

It is hard to overstate how common polysemy is. While not universal, it is, for example, extremely common in noun/verb pairs. Consider the following:

- (1) She drank a glass of *milk*.
- (2) She went out to the barn to *milk* the cow.
- (3) I am *staff* at the hotel.
- (4) The manager decided to completely *staff* the hotel with local teenagers.
- (5) You made quite the *gamble*.
- (6) He likes to *gamble* with other people's lives.

¹ There are different ways of cashing out "related". I cover this in 5.3.2.

(7) The politician put the motion on the *table*.

(8) The politician moved to *table* the motion.

Each of these examples represents a pairing between a noun and a verb denoting some action closely associated with the noun. Milk needs to be removed from the cow through a certain action – milking. A place of business needs to acquire staff – staffing. Gambling involves taking a risk – a gamble. These are not just a transposition of meaning from noun to verb or verb to noun (if such a transposition is even possible). The development of noun to verb or verb to noun involves a meaning shift, where the shifted meaning is metonymically related. While sentences (7) and (8) are used in identical contexts, “table” means slightly different things in both. “Table” in (8) alludes to placing a motion on a table, and historically owes its use to the noun form of “table” in (7), but the verb specifically refers to a formalized act in legislative bodies. Accordingly, there is little reason to think that these noun/verb pairs are anything but words with distinct but related meanings.

An interesting metasemantic puzzle arises, however, when we consider polysemous senses that share the same part of speech. Consider the following example from Vicente and Falkum (2017):

(9) John has his *mouth* full of food.

(10) Mary kissed him on the *mouth*.

(11) My *mouth* is sore.

(12) Watch your *mouth*.

(13) The *mouth* of the wine was dry.

(14) I have three *mouths* to feed.

(15) You can see the *mouth* of the river from here.

How are the semantic meanings of these uses of “mouth” related? In (9) “mouth” seems to refer to the oral cavity, “mouth” in (10) seems to refer to the lips, and “mouth” in (11) might be used to specifically discuss gums or the inner cheek. While it is plausible the first three uses all mean *oral cavity*, this does not explain the other four examples. Perhaps (12) and (14) still refer to the oral cavity via metaphor, but these uses of “mouth” are so well-worn it is plausible that even if they started their life as metaphors, they no longer are (Atkins & Rundell, 2008, 284). Similarly pushing against the metaphor picture, there is no word besides “mouth” for the opening of a river into a bigger body of water (although “estuary” and the Scottish-English “firth” covers a subset of mouths). Certainly “mouth” just is the name for mouths of rivers. Perhaps instead there is no metaphor here, and “mouth” shares a single, highly abstract meaning (Ruhl, 1989). This meaning would have to be more abstract than “opening” to capture (12), (13), and (14), and thus would likely either be so broad to include things that are intuitively not mouths or be so abstract to be unrecognizable to competent speakers of English.

In this chapter, I focus on how semantic externalists should theorize about polysemy. The overriding question here is: How should semantic externalists account for the phenomena where single strings of letters can have multiple separable and established patterns of use, and those patterns are related historically, in meaning, or in the minds of competent speakers of the language? I argue that semantic externalists should treat all established patterns of polysemous use – e.g., (9) to (15) above – as unique, if related, literal semantic meanings of a word. Call this view MEANING POLYSEMY. I argue for Meaning Polysemy in two ways. First, in Section 5.2 I argue Meaning Polysemy should not be a controversial position since established polysemy is the combination of branching meanings and related meanings, and the two phenomena on their own uncontroversially result in distinct literal meanings on semantic externalist frameworks. Pending some reason to think polysemy is special, we should similarly think established patterns of polysemy similarly represent

distinct literal meanings. Second, in Sections 5.4, 5.5, and 5.6 I consider other ways of accounting for polysemy, specifically as phenomena related merely to use, belief, or implicature. I argue that the first two views collapse into the third, but implicature does not adequately account for established polysemous uses of a word.

5.2 Introducing Meaning Polysemy

The thesis defended in this chapter is that semantic externalists, regardless of what they think ground semantic facts, should accept that established polysemes represent distinct literal semantic meanings of a word. I call this view MEANING POLYSEMY. The account I defend says that words have many meanings. These meanings are literal and grounded in extra-personal linguistic facts. In some sentences, “mouth” literally means the oral cavity. In other sentences, “mouth” literally means features of a wine, and in other sentences still “mouth” literally means the opening of a river. In this way, polysemous meanings are just a type of ambiguity with some etymological and psychological baggage that homonyms do not have.

Meaning Polysemy is not a contextualist account of meaning, as illustrated by:²

- (16) The *stars* aligned and the restaurant received four *stars* from the critic.
- (17) The athlete won *gold*, and later found out from pawnbroker that the medal was not actually made out of *gold*.
- (18) I can’t believe Tom put his *mouth* in that dirty water at the river’s *mouth*.

In these cases, each word is used to mean something else despite no change in context. Meaning Polysemy is also distinct from the microlanguage account of meaning defended by Ludlow (2014). Ludlow is interested in how conversational practices can modulate word

² Hazlett (2007, 673) seems to make the mistake in assuming polysemy is contextual by offering “tall” and “rich” as examples of polysemy.

meaning as conversational participants want and require. The sorts of polysemy considered here represent distinct and established patterns of usage. Rather than something that is implicitly or explicitly negotiated as part of a specific conversation, polysemous meanings pre-exist before the conversation starts based (partially) on extra-mental facts.

There are at least four reasons to resist Meaning Polysemy. First, Meaning Polysemy appears to break Grice's (1993, 47-50) maxim of not multiplying senses beyond necessity. According to the maxim, we should favor an explanation of a word's meaning that favors fewer senses over an explanation that posits more senses. Appearances of multiple senses, the Gricean thought goes, should be explained by implicature whenever possible. This principle has been criticized by Devitt on the grounds that if taken too absolutely, it does not allow for linguistic innovation (2013). However, it is unclear how absolutely the maxim's defenders take it to hold (see Hazlett, 2007), and as I discuss in more detail in Section 5.3.1, it is an uncontroversial feature of language that words sometimes split into multiple meanings.

Second, Meaning Polysemy holds that most common English words are ambiguous. According to Meaning Polysemy, even innocuous sentences like the following can be true or false, depending on which polysemous meaning a word contributes to the sentence:

(19) Bachelors are unmarried men.

(20) The day lasted 24 hours.

In (19), the sentence is false if "bachelor" means a young male fur seal or both sexist and false if it means someone who holds a bachelors degree. To see one of the false readings of (20), compare the true reading of (20) with another true statement about September 10, 2020 in St Andrews, Scotland:

(21) The day lasted 13 hours and 8 minutes.

We can even build sentences where "day" has both meanings in the same sentence, even if the swap between the two feels abrupt to us as readers:

(22) Given the day lasted 13 hours and 8 minutes, the day had 11 hours and 52 minutes of twilight.

In practice, we will rarely notice the ambiguities in sentences like (19) or (20). Context clues and obviously correct interpretations will stop us from looking for additional senses. Moreover, in the case of (19) especially, one sense is so dominant it will automatically come to mind over other senses (see Giora, 2003). Nonetheless, even if such ambiguities go unnoticed, Meaning Polysemy says that the sentences in (19) and (20) have literal false readings.

Third, and related to the previous two points, on Meaning Polysemy terms like “water”, “star”, and “tiger” can literally refer to things or properties that are not natural kinds.³ (16) and (17) already offer examples of this, where one use of “star” and “gold” refers to a natural kind and one does not. Other examples include uses of “tiger” that literally mean a person who is fierce and energetic. Meaning Polysemy holds these represent distinct semantic meanings of a word even though other explanations are possible. I return to this starting in Section 5.4.

I do not take these three concerns to be serious problems for Meaning Polysemy. My account is semantically profligate because language is semantically profligate. Words have lots of meanings, and we should not try to paper over that by appealing to non-semantic mechanisms. Nonetheless, there are at least three options for semantic externalists to explain the patterns of polysemous usage with non-semantic mechanisms:

USE POLYSEMY claims that polysemous senses are merely different but related uses of words. This is the most theoretically lightweight explanation of polysemy, and according to Use Polysemy, polysemy is an artifact. While it looks like words have different meanings, the linguistic patterns giving the appearance of different meanings are nothing

³ While this chapter does not discuss empirical work on polysemy (due to constraints on time rather than a lack of interest), there is some empirical data that suggest natural kind terms are polysemous. See Machery and Seppälä (2011) for a polysemy-based explanation of the sorts of phenomena discussed by Reuter (2019).

more than different patterns of use.

BELIEF POLYSEMY posits polysemy as a sociological phenomenon about wide-spread beliefs about words widely held among the community. This adheres most closely to the classic Kripke-Putnam line on natural kinds. Below in Section 5.3.3 I discuss the difference between folk uses of “berry” – which includes raspberries, strawberries, and blueberries – and botanist use of “berry” – which excludes raspberries, strawberries, and blueberries, but *includes* lemons, cucumbers, and bananas. Belief Polysemy claims that false folk beliefs about berries (or “berry”) are widespread and drive disagreement over the status of bananas and raspberries, but there is only one meaning of “berry”. This story can be expanded to non-natural kinds, since deference is a common feature of all language, where we defer to other people how words should be used (Sawyer, 2020) (see, however, Wikforss (2007)). According then to Belief Polysemy, polysemy is driven by similar mistakes about what the extra-personal linguistic facts in a community are.

IMPLICATURE POLYSEMY puts the phenomenon of polysemy squarely into the realm of pragmatics. Here, there is only one meaning of a given word and apparent polysemous meanings are instead patterns of conversational implicature. Sure, a strawberry is not a berry, Implicature Polysemy argues, but we do not have a word in English that picks out a category that lumps together strawberries, raspberries, and grapes. The opening of a river is not actually a mouth, but we can nonetheless use the word to communicate about the mouths of rivers. On this view, polysemous uses are misused words used to implicate something other than their literal referents. Sentences that implicate in this way are usually literally false but nonetheless express something true.

Since there are multiple accounts of polysemy, my defense of Meaning Polysemy will be two-fold. For most of the rest of the chapter, I defend the plausibility of Meaning Polysemy as an explanation of established polysemous uses of words. I focus on developing the framework and showing why semantic externalists should not be resistant to the idea that

words have multiple related meanings. Then, I reject Use Polysemy, Belief Polysemy, and Implicature Polysemy as adequate metasemantic explanations of polysemy. I argue that Use Polysemy and Belief Polysemy collapse into Implicature Polysemy, but Implicature Polysemy cannot explain the independence different polysemous senses have from each other.

5.3 Master Argument: Polysemy is not special

As discussed in the introduction, the assumption that words are univocal seems to be a common assumption in contemporary philosophy of language, against the view that most words have multiple related meanings. In this section, I argue for Meaning Polysemy by arguing that polysemy is not special. Instead, polysemy is merely a combination of other mechanisms that uncontroversially involve differences in meaning. Polysemy is the byproduct of diachronic meaning change – specifically the often ignored phenomena of branching meaning – as well as the phenomena of closely-related meanings. Both processes produce clear examples of distinct semantic meanings, so barring some further reason to think polysemy is special, the combination of the two should also produce distinct meanings.

5.3.1 Branching Meaning

Some English words' meanings have been remarkably stable over time. Historical reconstructions of Proto-Indo-European, the precursor language to all languages in the Indo-European language family (which includes English, Latin, Hindi, Hittite, and many others), have found that words like “water”, “daughter”, “liver”, and “swine” have, with gradual phonetic and morphological shifts, retained their rough meaning for at least 4500 years (*Indo-European Lexicon: PIE Etyma and IE Reflexes*, n.d.). These words are the

exception, however, rather than the rule. Most words have changed their meaning at some point in their history. Examples of semantic shift are somewhat common in the philosophical literature. G. Evans (1973) discusses an apparent meaning shift in “Madagascar” from meaning mainland Africa to its current meaning. More recently, as discussed in the introduction “meat” (e.g., Sawyer, 2018, 2020) and “salad” (e.g., Cappelen, 2018; Dorr & Hawthorne, 2014) have become standard examples, where the words’ meanings are thought to become less restricted over time.

Discussions of meaning change by semantic externalists have tended to focus on a specific type of semantic shift where “meat” gradually and univocally shifts from one meaning to another. At time T1, “meat” meant one thing, and, due to the vagaries of language, at time t2 “meat” meant something else. Contrast this SEMANTIC DRIFT to BRANCHING MEANING, where a word gains a new meaning while retaining its old meaning. Instead of a 1 to 1 relationship from time T1 to T2, meanings of the word have a one to many or many to many relationship. Consider “staff”, which can mean a wooden rod or the employees of a company. Both meanings developed from an Old English name for a walking stick. Through a byzantine set of branching meanings developed over more than a millennium, we ended up with the two different meanings.⁴

If meanings cannot branch, then either one of the following sentences must be literally false, or “staff” must mean the same thing in both:

(23) Tired from a long day walking, the hiker made a make-shift *staff* to help him walk.

(24) The university pays many of its *staff* below minimum wage.

It is unclear how “staff” could have a unified meaning in these two statements. Not only is one a mass-noun and one a count-noun, but one is made of a single bit of wood and holds a person upright while the other is made of multiple bits of flesh and holds a

⁴ In contemporary English, “staff” as in “chief of staff”, “staff sergeant”, (a company’s support) “staff”, and (musical) “staff” also share this etymology.

university upright. The meanings are so different that someone trying to do etymology without a historical corpus would be just as apt to assume they have a distinct origin than a shared origin. Nonetheless, the employee-related meaning of staff developed from military use of the word (“Chief of Staff”), which had originally been a reference to the baton (the “staff”) someone held as sign of authority.

Similarly, retaining a univocal-account by denying the truth of (23) or (24) is similarly off the table. It would be blatantly uncharitable to proclaim either (23) or (24) as false in situations where a hiker uses a long piece of wood as a hiking pole or a university exploits labor, respectively. Beside charity, denying the literal truth of one of the above sentences would also lead to extremely widespread false semantic beliefs. It is hard to state how common branched meanings are in English. To name just three more pairs whose connection between meanings has been lost to time, “magazine” meaning the print media and “magazine” meaning the place where ammunition is stored, “country” meaning a rural area and “country” meaning the political entity, and “address” – the location of something – and “address” – a speech – are all pairs of meanings that branched during English’s history. If we limit ourselves to only one of these meanings, then pretty much every English speaker has false semantic beliefs about at least some apparently innocuous uses of these words.

These branchings are all old and, we may suppose for the sake of argument, gradual and unnoticed by speakers. Perhaps at no point did any speaker think anything like “‘staff’ already means rod, but now I will use it to communicate about the group of people who work for a company”. This is important to note for polysemy, since many polysemous uses of words first arose during our lifetime, and many polysemous uses clearly arose intentionally. Nonetheless, arguments for the difference in meaning in old and (we can suppose) unintentional splits like “staff” also hold for new and intentional branching meaning.

To see this, consider the recent branching of “tweet”, which has two main contemporary uses. First, a tweet is a post on the social media platform Twitter. Second, a tweet is the sound a bird makes. Before Twitter’s founding in 2006, a “tweet” was, according to the OED, fairly monosemous. “Tweet” was used either as an imitation for a short high-pitched bird call (similar to “meow” or “moo”), or later, the onomatopoeic name for such a call (similar to “a meow” or “a moo”) or the action of making such a call (“she meowed”, “it mooed”). Then, around 2006 when Twitter was founded, the company used the language of a short bird chirp to describe the action and output of making short posts on their service.⁵ Twitter purposefully branched off a new meaning of “tweet”, choosing the word “tweet” because of its associated meaning. This does not, however, mean the new use retained the existing meaning. Compare (23) and (24) to:

(25) The bird tweeted in disdain at its destroyed nest.

(26) The president tweeted in disdain about negative news coverage.

Again we are left with the option of accepting the meaning branched, or arguing that “tweeted” means the same thing in both sentences. In the latter case, either both sentences are literally true because they share a general meaning or one of the sentences is literally false (presumably (26)) because one of the uses does not correspond to the word’s meaning. One of these univocal-meaning responses is likely more attractive than with “staff”, but we should not confuse associated imagery with having the same meaning. Even though Twitter intended to associate the actions with birds, this is a non-semantic feature of (26). The association is psychological rather than semantic. Like the above statements about “staff”, “tweeted in disdain” in the two sentences literally describe two different sorts of actions. To my eyes, “tweet”, if it was ever a metaphor, is dead (Grey, 2000; Lakoff, 1987), judging by the infelicitousness of the following correction:

⁵ This was only one part of a package of imagery, as “twitter” has two distinct meanings of a bird sound on one hand and meaningless words on the other.

(27) The president tweeted in disdain about negative news coverage. Well not actually, the president posted on Twitter.

Compare (27) to a case where I am commenting on how complicated a paper is:

(28) I think this paper is relevant to my work, but it's written in Sanskrit. Well not actually, it just uses complicated hyperintensional logic.

Like “tweet”, similar pairings can be made with other recent and intended branched meanings such as “mouse” (computer device vs animal), “cloud” (server-based storage vs collection of water vapor), and “string” (fundamental entity in physics vs component of fabric). If we deny that language can split and shift in this way, we have to admit that words with a shared etymology mean the same thing.⁶ This entails extremely widespread false beliefs about language – either (23), (24), or both means something dramatically different than we would otherwise think – and this eliminates all linguistic innovation that is not either the product of the introduction of a new word or phrase or the gradual shift of a single unified meaning.

5.3.2 Related Meanings

For the purposes of polysemy, there are two ways we can cash out related meanings, which correspond to different ways of cashing out the notion of polysemy. Either polysemous sense are related cognitively or psychologically in the minds of the users of a language, or polysemous senses are related in a non-psychological semantic or linguistic way. Polysemy of the former notion, what we might call COGNITIVE POLYSEMY, has been the focus of cognitive linguists, such as Lakoff and Johnson (1980). When words are polysemous, they share some sort of shared mental or cognitive representation (what exactly that representation is differs depending on the account) that is not shared between homonyms.

⁶ Something like this happens when a certain breed of philosophers looks to a word's etymology to support their account. This is a well-known fallacy in linguistic circles.

In contrast, what we might call LINGUISTIC POLYSEMY occurs when one word has multiple use connected by linguistic or semantic properties, such as coextension or etymology. This notion of polysemy is common in fields like historical lexicography, where different senses are grouped by headword, which are chosen based on etymological connection.

Note that words can be connected in either, none, or both of these ways. As discussed above, words like “staff” and “magazine” have historical linguistic connections, but their different senses are likely not lumped together by many English speakers. At the same time, words can also be jointly represented without sharing an etymology. For example, some English speakers likely think “post” (a job) and “post” (a piece of wood in the ground) are connected due to sentences like

(29) The guard stood at his post all evening.

However, the job-use comes from French *poste* and Italian *posto* in the 16th century and the support-use was present in Old English and likely came from a combination of Latin *postis* and French *post*.

Whether we think about the relevant sense of related meanings as cognitive or linguistic, non-polysemous examples of both are clear examples of distinct meanings. Words can be largely co-extensional, have related intensions, or be connected in the minds of speakers without meaning the same thing. To run is not the same thing as to sprint, despite both involving moving on two legs at an increased speed. Cats are not dogs, despite being connected in many people’s heads as prototypical house pets. To be arrogant is not the same thing as having hubris, even though many people will be hard-pressed to distinguish the two. One of semantic externalism’s strengths is the ease by which it can explain why closely-related words like “hubris” and “arrogance” cannot be distinguished by many, if not most, people but nonetheless have distinct meanings. It does not matter if the average person can distinguish the word’s meanings because the meanings of the words are not (solely) grounded on their mental states. Related meanings of different

words are different meanings.

5.3.3 Polysemy: Related and Branched Meanings

So far I have defended that branched meanings (Section 5.3.1) and related meanings (Section 5.3.2) both result in numerically distinct meanings. Polysemy is just the combination of both phenomena, when one phonological or morphological unit develops multiple closely related meanings through branching. Both processes lead to distinct meanings of a word separately, and they lead to distinct meanings of a word together. To see this in practice, consider the following exchange from the quiz show QI. The host, Stephen Fry, makes an innocuous request:⁷

Stephen Fry: (To contestants) What I want you to do now is name a berry.

Alan Davies: Blackberry, strawberry, raspberry, blue-

[A klaxon sounds indicating Alan Davies has said something “obvious but wrong”]

Stephen Fry: [After the klaxon] Hold on a minute because we have a lot of catching up to do. Blackberry, strawberry, raspberry...

Clive Anderson: Isn't a banana?

Stephen Fry: A banana! I give you good points for that. It is a berry, yeah.

[Later]

Clive Anderson: But why is a banana a berry?

Stephen Fry: A berry is a fruit that *contains* stones, pips, things like that.

There must be more than one [stone, pip, etc.] (*Cleve Crudgington*, 2005)

⁷ Thanks to Fenner Tanswell for pointing out that this sort of exchange is regular in QI.

Alan Davies is chastised for saying a blackberry is a berry, while Clive Anderson is rewarded for saying a banana is a berry because, according to Stephen Fry, a banana is a berry and blackberries are not. According to the standard univocal assumption of language, since botanists tell us berries are fruits containing stones, pips, etc., “berry” only ever literally refers to those things. It may non-literally refer to something else as slang or people might have false beliefs about berry-hood, but “berry” only has one meaning. In the above exchange, Stephen Fry can be seen playing the role of the univocal theorist. Even though it is commonly thought that blueberries, strawberries, and raspberries are berries, they are not, full stop.

Later in the exchange, Clive Anderson has a change of heart:

Clive Anderson: On what sort of bizarre definition of the word “berry” does it mean –

Stephen Fry: That of people who study fruits

Clive Anderson: I’m afraid we the speakers of English –

Stephen Fry: I know, but you know this is QI! (*Cleve Crudgington*, 2005)

Even Clive Anderson, someone who was just rewarded for saying that bananas are berries, is torn about his initial answers’ relationship to English. The reader should be torn too. A berry parfait should not include eggplants, cucumbers, and gourds – all berries according to the above definition. Similarly, a dish of berries and cream could include strawberries, raspberries, and blueberries, but according to the above definition, they are not berries. What has happened is that, like “staff”, “berry” branched. Around 1809, botanists started using “berry” to specifically talk about fruits with certain botanical features, but this was different than the preexisting and broader use of “berry”. This preexisting use comes down to us today in the form of the everyday use of “berry” we use when ordering fancy iced coffees or describing to a guest what smoothies we can make for

them. While 1809 may seem relatively recent linguistically, the botanists' sense of berry is actually a few decades *older* than the employee-sense of "staff", which the OED dates to around 1837.

The two uses of "berry" have closely related meanings while "staff" does not, and this helps enable Stephen Fry's pedantic comments. The two uses of "berry" are both about small fruits. Nonetheless, the two uses have a surprising lack of overlap. The overlap between the two uses of berries is, with the exception of the grape, limited to a few less-common fruits like gooseberries and the berries of mistletoes.

Because botanists intend "berry" to be a natural kind term, "berry" has an extra level of complexity not shared by polysemous words like "mouth". Perhaps people are wrong to think strawberries are members of the natural kind berries. I consider and reject this sort of response in relationship to Use Polysemy and Belief Polysemy in Sections 5.4 and 5.5 below, but for now it is worth noting that even if one sense refers to a natural kind, not all senses have to. Consider "string". If string theory is correct, strings are a natural kind. That does not mean, however, that that all uses of the count noun "string" refer to natural kinds. Otherwise the following sentences would not have the literal meaning they seem to have:

(30) The strings on the violin snapped.

(31) The dog ate a whole string of sausages.

(32) After a summer of intense training, Jamie made it on to the football team's second string.

(33) Amir went all the way to the crafts store to buy a single string.

Moreover, only a small portion of polysemous words have a use related to a natural kind. So even if "berry" only means what botanists tell us it means, this leaves words like:

- “window”: the space on my computer screen/hole in the side of a building/the place in a train station someone buys tickets
- “vice”: a tool that holds stuff/certain unsavory habits
- “bank”: A building that holds money/the corporation that holds money
- “film”: The physical medium old movies are stored on/the movies stored on the medium
- “trolley”: The vehicle obsessed over by ethicists/British English for shopping cart

Something needs to be said about the relationship between these distinct uses. Since they are the combination of the otherwise unremarkable phenomena of branching meanings and related meanings, we should then, at least without further reason to think they are special, assume they also have distinct meanings. In this way, like “bank” and “bank”, “window” and “window” are, semantically, cases of run-of-the-mill ambiguity. Consider:

- (34) Always looking to spend time near the sound of running water, I often pass *banks* during walks.

Bank, on its own, is ambiguous between two different homonyms. However, context helps us come to the correct interpretation of the sentence, and we come to know that the person enjoys walks near the sides of rivers. This is distinct, however, from saying context *determines* the meaning, of “bank” here. Rather, faced with the issue of figuring out what exactly the semantic and/or speaker meaning of the sentence was, we need to rely on information beyond the morphology or phonology of the utterance. Similarly, when faced with words with meanings that are etymologically or psychologically related, the process is no different.

- (35) Unsure of what the company’s hiring practices are like, Raz reached out to former *staff*.

(36) Xiao bought a Netflix subscription but was disappointed by the quality of available *films*.

The key difference between “bank” and “films” is not that they are different in semantic kind. Rather “films” has psychological baggage, where, for contingent reasons, the different senses are closely represented cognitively (Klein & Murphy, 2001; Vicente & Falkum, 2017). But semantic externalism is not beholden to psychology in the same way as semantic internalism, so if meaning ain’t in the head, our propensity to associate meanings should not in itself have an effect on the semantics of words.

This section has argued the view that polysemous words represent distinct literal meanings – that is, Meaning Polysemy – should be the default view of polysemy. I argued for this by comparing polysemy to branching meanings and related meanings. Branching uses of a word, such as happened to “staff”, “country”, and “address”, lead to multiple meanings of the words. Similarly, words with similar meanings, such as the subtle difference between arrogance and hubris or a sofa and a loveseat, result in different meanings. When brought together in the form of polysemous senses we should also assume the different established uses correspond to different meanings. The rest of the chapter looks at other metasemantic explanations for polysemy and finds them lacking. In order, I consider and reject explaining polysemy on the basis of merely different uses, different beliefs, and finally different implicated content. Use Polysemy collapses into Belief Polysemy or Implicature Polysemy because *something* needs to explain the differences in uses. Belief Polysemy does not explain polysemy because polysemy is not correctable in the same way as other metalinguistic mistakes. Finally, Implicature Polysemy both misunderstands the way in which polysemous meanings are established and falsely predicts understanding of a “primary” sense of a word is key to understanding what is communicated by a non-primary polysemous sense.

5.4 Use Polysemy

The first possibility is accounting for polysemy as a phenomena of use. Use Polysemy holds that any apparent polysemy reduces to different patterns of use and nothing more. We can find the heart of such a response in Abbott (1997). Abbott is considering cases where substances like the contents of the Great Salt Lake in Utah is called “water” whereas substances that have purer amounts of H₂O, like beer and soda, are not called “water”.⁸ This raises an issue for traditional externalist accounts because it suggests “water” is not tracking the natural kind H₂O, but descriptive features such as being wet, liquid at room temperature, tasteless, etc.

Abbott’s defense of the identity of water and H₂O in light of these patterns of use is straightforward. Consider that I do not regularly call myself a mammal, and I cannot think of the last time I have been called a mammal. Nonetheless, I am a mammal. More esoterically, I can say with certainty I have never been called a bilaterian nor have I heard anyone ever called a bilaterian, but I am a member of the evolutionary category of animals with bilateral symmetry. Similarly with water, there is a gap between language and metaphysics, and things can belong to a category even if people do not call it a member of that category.

Pausing Abbott’s argument, an adjustment to this observation can provide an argument for Use Polysemy. Abbot is rejecting the claim that being called “P” is a necessary condition for being P. This is why beer is water even if people do not call it “water”. In contrast, in order to limit the amount of uses that correspond to the semantic meaning of a word, Use Polysemy needs to reject that being called “P” is a sufficient condition for being P. Along with the additional claim that polysemous words only have one semantic meaning, denying the sufficiency condition allows a picture where words have a single meaning and any use that deviates from it is merely a pattern of use. On this view,

⁸ See also Bloom (2007), Pietroski (2018).

taking these patterns of use as semantic meaning is taking language use too seriously; calling a part of a river a “mouth” does not make it an actual mouth.

In Use Polysemy’s favor, the sufficiency condition *is* almost certainly false. Non-literal speech means that being called “P” is not sufficient for being P. Juliet is not the sun despite Romeo’s assertion otherwise. Resuming Abbott’s argument, however, the problems for Use Polysemy become evident when we consider why there is a gap between being called “P” and being a P. Abbott explains the disconnect between terms and corresponding predicates by appealing to pragmatics. According to her, we do or do not call water “water” depending on what we want. So, on her account, while “beer” is water in the same way that distilled spring water or the stuff that fills the Great Salt Lake is water, I ask for “beer”, she argues, and not “carbonated water containing fermented hops and grains” because in context it is easier to use “beer” to get my point across.

Use Polysemy needs to tell a similar story to explain why particular uses persist despite a single meaning. Unless the distinct patterns of usage are either a cosmically unlikely coincidence or illusory, there is some mechanism driving the different uses of a word. When I and millions of other people call a part of a river its “mouth”, either I falsely believe “mouth” literally means the mouth of a river (and Belief Polysemy is true), I intend for “mouth” to communicate something about the opening of the river (and Implicature Polysemy is true), or I truly believe “mouth” literally means the opening of the river (and Meaning Polysemy is true). Barring some argument that there is in fact only one use of any purportedly polysemous word, Use Polysemy therefore collapses into one of the other views considered here.

5.5 Belief Polysemy

Belief Polysemy is the standard externalist account of instances where use of a word diverges from scientific facts or knowledge (more on this in Section 7.2.2). On this picture,

people fail to use “water” to talk exclusively about H₂O, call a disease in their thigh “arthritis”, or call a realistic tiger-like robot a “tiger” because they have false beliefs about the world. While these classic examples focus on object-level beliefs, more recently the verbal dispute literature (Belleri, 2018; Chalmers, 2011; Inga, 2018; Jenkins, 2014) has focused on differences in uses based on beliefs about language, where disagreements are partly or entirely caused by different people having different beliefs about the meaning of a word or phrase.

The examples used in both semantic/content externalist literature and the verbal dispute literature usually revolve around words thought to be univocal. Chalmers (2011), for example, initially considers verbal disputes where the verbal dispute causes parties to express compatible propositions by apparently incompatible sentences. Where one person says “free will is not compatible with determinism” and the other says “free will is compatible with determinism”, rather than expressing proposition P and not-P, the sentences are actually expressing propositions P and Q. Shifting then with an eye towards semantic externalism, Chalmers changes the set-up of verbal disputes so that each party is expressing incompatible propositions as well, saying to do so, “we need only tweak the dispute as above so that the two parties use the key terms with deference to a linguistic community in which those terms are used univocally.” So insofar as “free will” is a verbal dispute in this way, Chalmers couches his discussion in the assumption that it is a verbal dispute caused by incorrect beliefs about the semantics or metasemantics of “free will”.

Drawing from these sorts of examples, Belief Polysemy claims that polysemous uses of a word is driven either by false belief about the world or false belief about language. For all the examples of polysemy discussed above – window, mouth, vice, film, trolley, tweet, string, berry, star, gold – these only have one meaning despite people’s beliefs they have other meanings. This view already faces one problem raised in relationship to “staff” in Section 5.3.1, where because of the extent in which people use polysemous words in

different ways means that Belief Polysemy entails false beliefs are extremely widespread. Setting this worry aside, Belief Polysemy entails that certain uses of a polysemous word should be correctable. But this is not what we see. The sort of corrections that solve the sorts of false beliefs evident in verbal disputes or mistaken beliefs about natural kinds feel infelicitous when applied to many examples of polysemy. Consider the following examples of corrections about water and “water”:

Fred: I was walking down by the river and saw the water was really mucky.

Francesca: Ah, because of all the dumping, the river is actually just chemical waste now – I think mostly some strange liquid form of lead.

This is a correction about the world.

Fred: I can’t believe they discovered water on Twin Earth with this weird chemical makeup. What was it? XYZ?

Francesca: Actually, it’s not water since it is not made of H₂O.

This is a correction about the nature of water.

Fred: It’s cool that “water” is at least 6000 years old and has only ever been used to talk about H₂O.

Francesca: Historically, British alchemists also called aqua fortis “water”.

This is a correction about “water”.

All of these are good-faith attempts to correct different mistakes involving water and “water”. Consider instead similar examples built around “mouth”. Many polysemous words seem to have a primary sense (Recanati, 2017), and “mouth” is no different. If “mouth” has a singular meaning, it is the oral cavity. Treating *oral cavity* as the proper meaning of “mouth”, consider the following:

Jack: When I was in grad school, I had three mouths to feed and almost no income.

Jill: You only have one mouth.

This is a correction about the world

Jack: When I dropped the pickle jar, the mouth broke clean off.

Jill: Mouths are the oral cavities of people, not parts of jars.

This is a correction about what a mouth is

Jack: It's neat that "mouth" is so flexible. It can mean a part of a rivers, a person's nutritional needs, a feature of wine, and even just the oral cavity.

Jill: Actually "mouth" *only* means oral cavity.

This is a correction about "mouth".

If polysemous uses were just false beliefs about a univocal meaning of a word, Jack and Jill's discussion of mouths should line up with Fred and Francesca's discussion of water. In each exchange, Belief Polysemy holds, a false belief is expressed and corrected. Instead, each of the exchanges involving "mouth", while parallel to the exchanges about "water", are – if not outright hostile – uncharitable to Jack in a way Francesca's corrections about water are not uncharitable to Fred.

Here we can see that Belief Polysemy faces a similar challenge as Use Polysemy. Jill's responses are so hostile because Jack is successfully communicating or saying something, but what Jack is communicating or saying is not about oral cavities. This is why Jill's response seems so hostile. We know what Jack is communicating when he says "the mouth [of the jar] broke clean off", and it looks like Jill's failure is either the failure to recognize this intention to communicate something about the pickle jar or a failure to understand the meaning(s) of mouth. So either, Jill is failing to pick out implicature *or* Jack and Jill

are engaged in a verbal dispute that Jill is wrong about. In the former case, polysemous uses of a word are explained by implicature, in the latter case, polysemy is a feature of meaning. Since I have already made my case for polysemy being a semantic feature of language, we will turn our attention to the possibility of explaining polysemy in terms of implicature.

5.6 Implicature Polysemy

Building from the discussion and problems with Belief Polysemy, Implicature Polysemy also holds that words have a single literal semantic meaning. Instead of explaining patterns of usage as mistakes, Implicature Polysemy treats the use as intentional. Sometimes speakers purposefully misuse the literal semantic meaning in order to pragmatically communicate something else. Returning to example (14), “mouths to feed” is literally a category mistake. An oral cavity cannot be fed. But, by saying “I had three mouths to feed”, the speaker is communicating something true, even though what the speaker said is literally false.

For present purposes of exploring whether polysemy can be explained by Implicature Polysemy, we need to be careful about which uses of a word we consider. There are certain metonymic uses of a word that I do not want to deny involve implicature. Instead, I will reject the claim that *established* patterns of polysemy involve implicature.

Many polysemous senses have at least two distinct stages of life (see discussion of “swallow” in Recanati (2017, 2019)). First, some words are used in a novel and creative way to express something that the word does not typically express as a form of linguistic innovation. Sometimes they are short-lived, but sometimes the innovation propagates, the novel use ceases being novel and becomes standard linguistic practice. During the sense’s start of life at the stage of innovation, implicature may be involved. When polysemous uses are introduced, the new uses of a word do not have a semantic meaning to express

what is intended, nonetheless it manages to communicate something. E.g.,

(37) Discovering that “wombat” is monosemous, Hussiny set out to *polysemy* it.

(38) Because of COVID-19, the university hired *sanitizers* to do rounds between classrooms.

To my knowledge, I have never heard “polysemy” used as a verb or “sanitizer” used to refer to people, yet it should be clear what is being communicated by both (37) and (38). Metaphysically, because there is no pattern of language or other relevant metasemantic fact, there’s no corresponding meaning for “wombat” or “sanitizer” to have at the moment of utterance. Psychologically, because the use deviates from the person’s own vocabulary, the listener plausibly has to go through a mind-reading process to figure out what the speaker’s (my) intention was in that situation, and if the speaker is aware that they are using the word in a novel way, then there’s the corresponding intention for the listener to pick up on the speaker’s intention. Therefore, it is not surprising if Implicature Polysemy explains lexical innovation such as novel metonyms. The issue currently at play is whether Implicature Polysemy captures polysemous use after the polysemous use has become widespread among the population.

Implicature Polysemy is not a plausible account of established polysemy. First, as discussed extensively above, many polysemous uses of a word have a long history of use. “Berry” in the botanist sense dates back to at least 1809, “mouth” as the person who eats has a first recorded use in 1553, and use of “swallow” to mean to engulf (as in “an ATM swallowed/engulfed my card”) dates to the late 12th century.⁹ These are already established patterns of usage and will be known by a competent user of En-

⁹ I sometimes see philosophers discuss cases like this (see Recanati (2017) with ATM “swallowing” cards and Maitra (2018) with both “stalking” people online as well as “harassment” when modified into “sexual harassment” as instances where a word’s meaning is expanded to include new sorts of cases. I admittedly cannot make sense of this sort of claim. It seems to me that the intension/dictionary meaning of the word has remained the same, but we update our beliefs regarding the term’s extension. This then causes a change in our use.

glish. Therefore, the metaphysics and psychology look more like regular semantic meaning than implicature. Any interpretation of speaker's intention does not require any sort of mind-reading, or at least no more mind-reading than determining which of competing homonyms/homographs a speaker uses. From the point of view of the listener's understanding, polysemy thus presents the listener not with a problem of inferring non-literal use but disambiguating between alternative senses of a word.

Second, this account requires the word means *something* in order to infer implicatures from. This leads to trouble for Implicature Polysemy. In phrases like "swallowing my card", it is clear what the primary meaning supposedly is – the physical act of swallowing food. Similarly, the best candidate for primary meaning of "mouth" is the oral cavity. But consider: "square", the tool used by carpenters to draw 90 degree angles, predates "square" the geometric entity. Does that mean the primary meaning of "square" is the tool? Perhaps historical linkage is not important, but consider "bank". Is the primary meaning the building the money is stored in ("I walked to the bank to deposit the check") or the company that owns the building, ("The government bailed out the bank")? In these cases, it is not obvious what the primary meaning is. Similar issues arise with "state" (having a property vs the political entity), "country" (rural area vs political entity), "lot" (e.g., "a lot of stuff" vs "the lot up for auction"), or "instrument" (music device vs more general task-focused apparatus). While there are general features of polysemy that lead to regular shifts from concrete meanings to abstract meanings (Lakoff, 1987), this is not true of all polysemy, and it is unclear how Implicature Polysemy can give a clear account of which meanings are the primary meanings being used as a vehicle of non-literal implicature for all other polysemous uses.

Third, and most problematically for the Gricean picture of polysemy, even when a word has a plausible candidate for primary meaning, understanding primary meaning is not necessary for understanding the purportedly derivative, non-literal meaning. Poly-

semous senses are able to stand on their own, and using or understanding a sense does not require using or understanding other senses. To illustrate this, imagine a non-native English speaker was learning English, and in her original language, “mouth” was not polysemous as it is in English.¹⁰ Imagine further her English textbook’s chapter for anatomy was missing, but the chapter for geography was not. The non-native speaker could plausibly learn how to use “mouth” in relation to the part of a river, competently deploy “mouth” in sentences, understand the propositional or truth-conditional content of sentences using “mouth” in this way, use it with the intention to use it as native speakers do in relation to a part of river, etc., Name any test for semantic competence you want, an English language learner could be competent in “mouth” regarding rivers without knowing the word’s relationship to anatomy. Positing a step of implicature does no work beyond maintaining the claim that polysemous words only have one meaning.

Implicature Polysemy suffers from three setbacks. First, established polysemous senses look both metaphysically and psychologically like their own meanings. Second, polysemous senses do not always have clear primary meanings. Instead, some words have multiple plausible candidates for the “real” semantic meaning implicatures are derived from. Third, even when there are clear primary meanings, the primary meanings are not necessary for understanding the word’s use. In a sense, the second and third problem are consequences of the first. As I defended in Section 5.2, established polysemous senses have everything in place to be their own literal meaning. We should not overthink polysemy. There are tricky questions to still be asked about the exact content and delineation between distinct senses, and I invite the reader to look up the OED entries for “water” and “good” for a taste of this challenge – although these may ultimately fall under the purview of linguists. Nonetheless, we should be ready and willing to admit that words, especially high-frequency words, have multiple related literal meanings. This is also true,

¹⁰ There is some question of how universal polysemous meanings are. See Lakoff and Johnson (1980); Vicente and Falkum (2017).

as we will see in the next chapter, in technical settings such as philosophy.

5.7 Conclusion

In this chapter I argued that semantic externalism should claim that polysemous uses of a word typically track distinct meanings of that word. Polysemy is a well-known phenomena among linguists, who take most words to have multiple related meanings. Semantic externalists, I have argued, have no plausible account of polysemy except that different senses track different semantic meanings of the word. Accounting for polysemy as mere related uses has to claim implausibly that no mechanism drives the apparently distinct uses. Accounting for polysemy as mistaken beliefs is uncharitable to users of polysemy because it does not account for the way they intend for polysemous uses to communicate certain things. Accounting for polysemy in terms of implicature fails to account for the established patterns of polysemous senses and the way in which no inference seems to be necessary to understand what a user of a polysemous sense is communicating (at least no mind-reading over and above that is needed for homonyms.)

Instead, semantic externalists should hold that polysemous words have polysemous meanings. The basis of this claim is straightforward. Sometimes words develop two meanings from a single meaning, such as happened with “staff”. Sometimes words mean the same thing as other words, as is the case with “napkin” and “handkerchief”, but the meanings are still distinct. Polysemy is the co-occurrence of semantic branching and similar meaning and also produces numerically distinct meanings. In this way, polysemy is a form of ambiguity like homonymy.

Not everything that can be said about the metasemantics of polysemy has been said here. In particular, I have not waded into the literature on conventional implicature (Potts, 2007) and considered whether “mouth” works like Griceans take “but”, “still”, “or”, or – when communicating temporal or causal connection – “and”, to work. This

omission is in part due to time. It is also in part due to my belief for reasons outlined here and in the next chapters that Griceans are wrong to hold that one of the main sets of words thought to have conventional implicature, conjunctions in English like “and”, “or”, and “but”, are monosemous whose literal meaning lines up with their corresponding logical connectives. Omissions like this are also in a large part due to my goals. While I think polysemy can and should be theorized by philosophers of language as run-of-the-mill ambiguity, my larger goal in the dissertation is to offer a plausible framework for understanding semantic meaning in an externalist framework which would allow the use of lexicographical methods by metaphilosophers. In the aftermath of Kripke (1981), it is easy to think, as many metaphilosophers do, that meaning for many words, including words like “knowledge”, “justice”, “good”, “right”, “intuition”, is something for philosophers to find by hunting essences. And, while I do not want to deny that we can find associated essences for many of these words, we cannot assume the simplistic picture that words and meanings have a 1 to 1 relationship. The truth is more complicated, including, as I discuss next, in technical settings such as philosophy.

Chapter 6

Jargon

In the previous chapter, I defended the claim that even on semantic externalist accounts, polysemous uses of words often represent distinct semantic meanings of that word. Semantically, polysemy is just a type of ambiguity where the meanings carry a close historical or psychological connection. In particular, polysemy is the combination of branching meanings and related meanings, both of which are semantically unremarkable.

Polysemy is important to the epistemology of philosophy because “intuition” is polysemous. This polysemy has driven confusion among metaphilosophers who have assumed that the intuition-debate is about some unified overarching category of mental states instead of multiple related categories. However, the polysemy of “intuition” differs in important ways from the examples discussed in the previous chapter. Many polysemous senses of “intuition” only appear in philosophical settings among a small number of metaphilosophers. The senses are not a standard parts of English-speakers’ vocabulary, nor have they spread beyond a certain small subcommunity of speakers. Instead, the senses were developed to serve the theoretical and communicative interests of a handful of specialists.

This chapter explores the interaction of polysemy and technical language, or jargon. Specifically, I discuss why existing methods for spotting jargon that have come out of sci-

ence and healthcare communication research are not fit for purpose in identifying whether or not “intuition” has multiple related senses among metaphilosophers. In response, I propose a number of methods specifically designed to test for whether a word has polysemous meanings specific to technical settings such as philosophy.

In Section 6.1, I do some stage setting and discuss extant research on jargon. I then move to specific methods that have been developed to identify jargon in 6.2. I highlight how these tests are not fit for purpose in cases where standard English words have been co-opted by specialists and used in a technical way. In Section 6.3, I explore the strengths and weaknesses of a number of tests, including the anaphor and zeugma tests (6.3.1), certain methods of corpus linguistics (6.3.2), and the presence of sense-talk (6.3.3).

In the next chapter, I apply the lessons here to the intuition debate and argue that “intuition” is well on its way to becoming polysemous in the mouths of metaphilosophers. I argue this by borrowing methods from lexicography to study how philosophers theorizing about intuitions have used “intuition”. The lessons from that will help clarify how the intuition debate should continue going forward. It will also highlight the importance of lexicographic data to the philosophical study of philosophy.

6.1 Introducing Jargon

This is a chapter on jargon. Unfortunately, “jargon” is polysemous. The word “jargon”, especially in British English, is often used broadly and derogatorily to mean any sort of unintelligible or overly complicated speech. I will be avoiding that usage here. Instead I intend “jargon” to be understood by another dictionary definition: *the vocabulary unique to a particular group of experts*. I am particularly interested here in the subset of such language that serves specialized and teleological ends. These are words or meanings of words used by specialists for the purposes of communicating about that specialization where the words or meanings have limited use outside of that specialization.

On this understanding of jargon, there is no hard and fast distinction between jargon and non-jargon, especially over time. Words that start their life in a technical community can leak out into wider usage for a variety of reasons. Impenetrable acronym-laden classificatory names that would otherwise never leave the pages of scientific journals, such as “COVID-19”, leap into public usage when work in scientific settings becomes relevant to day-to-day life. At some point in spring 2020, “COVID-19” became standard English, following the path from jargon to standard English also taken by “DNA”, “microaggression”, and “radiation”. When non-specialists use words that have entered the everyday lexicon, they lack subject-specific expertise on what they are talking about. Consequently, sometimes the word’s meaning is retained, but other times the cross-over results in a new polysemous sense. For example, internet memes are only a subset of the memes that meme-theorists study.¹ Note that the lexical relationship between expert vocabulary and non-expert vocabulary does not only run in one direction. Terms that are common English may fall out of general usage but retain specialized use, or, to anticipate discussion in the next chapter, a standard English word can gain a specialized polysemous meaning.

Over the last decade, the fields of healthcare communication and science communication have begun studying jargon in healthcare and science. In both healthcare and science, jargon is common in communication (e.g., Castro, Wilson, Wang, & Schillinger, 2007; Links et al., 2019; Subramaniam et al., 2017). At the same time, technical ideas in healthcare and science need to be communicated to non-experts. This can involve doctors and nurses speaking to patients about their health or scientists explaining findings to regulators. Because the information being communicated to non-specialists is often complicated or arcane, jargon is often key for a speaker to express what exactly it is they are trying to communicate. Jargon can offer expressive power not found in everyday language, where not using it would force specialists to use cumbersome or inaccurate

¹ In fact, there is reason to think this shift in meaning happened with “COVID-19” and “COVID”. The name was coined to discuss the virus, but now it is used to discuss the disease as well. See Landes (ms).

descriptions.

One of the more commonsensical problems with jargon is that since jargon is not understood by non-specialists, jargon hinders non-specialists' comprehension. On one experimental estimate, for readers to adequately comprehend a relatively uncomplicated text, they need to understand 98% of the words in the text (Hu & Nation, 2000). Otherwise, readers fail to comprehend the points being expressed by the work. In healthcare and science communication, where texts and speech cannot by the nature of the setting be uncomplicated or low-stakes, observational studies have found striking results. In one study involving jargon by chemotherapy nurses, jargon use by nurses led patients to be confused about things as basic as the number of treatments they would receive (Schnitzler et al., 2017).

Even when jargon terms are defined in texts, they are still difficult for non-specialists to understand (Bullock, Colón Amill, Shulman, & Dixon, 2019). Difficulty of this form leads to striking downstream effects on how communication is consumed. It takes cognitive effort to apply a novel definition to an unfamiliar term over the course of reading a work. The increased effort can then bring biases online and lower feelings of knowledge. This causes readers to find whatever is written less persuasive – even if they understand what is being communicated (Bullock et al., 2019; Shulman, Dixon, Bullock, & Colón Amill, 2020).

Given the technical nature of philosophy, jargon is undoubtedly a barrier to communication between philosophers and non-philosophers. In the process of developing a better understanding of philosophical issues, philosophers have developed language special to the craft of philosophy, and regular people do not know what these words mean. We should not, however, think this barrier is only limited to communication between philosophers and non-philosophers. Long-gone are the days when a sufficiently well-read reader could be on top of every major contemporary philosophical debate. Given the explosion in the

number of philosophers and the amount of work in philosophy over the last century or so, work in philosophy is often insular. Someone working on Kant cannot be expected to understand the vocabulary of contemporary analytic epistemology or vice versa. Philosophy does not have one single vocabulary shared by all practitioners, and the same mechanisms found to reduce the persuasiveness of jargon-laden text in healthcare and science communication (namely, process fluency) has been found to have an effect on how people react to philosophical texts (Weinberg, Alexander, Gonnerman, & Reuter, 2012).

Complicating the picture, the barriers to philosophers' understanding other philosophers' jargon do not necessarily track the distinction between subfields. While subfields (such as epistemology or normative ethics) do often have their own vocabulary, sometimes specific debates within those subfields do too. At the same time, theoretical frameworks or traditions can use distinct vocabularies, even *across* different subfields. Virtue epistemologists do not talk about testimony in the same way as non-virtue epistemologists. Similarly, naturalists do not talk the same way as non-naturalists – regardless of the subfield.

Here then is the first lesson we can draw about jargon:

FIRST LESSON OF JARGON: Jargon can belong to specific groups whose boundaries do not cleanly line up with disciplinary or sub-disciplinary boundaries.

Notice the similarity between jargon and slang, since circulation of slang is by definition limited to social boundaries. Stereotypical slang does share a number of features with jargon. As with jargon, slang splits existing linguistic communities between the initiated and the uninitiated, and like jargon, slang is heavily dependent on co-opting existing words or developing neologisms that combine well-known roots and words to new ends. Indeed, jargon can be thought of as a proper subset of slang, and much of what is said in this chapter is also applicable to slang. My present goal, however, is to think about

words with a particular teleology, namely, words that are developed to expand the expressive power of a language in order to aid communication around some specific craft or knowledge-producing discipline. To this end, “stan” and “finna”, while limited to a social subgroup, are merely slang and not jargon.

6.2 Existing Approaches to Identifying Jargon

Now that I have raised problems associated with jargon and made some clarifying remarks, the rest of this chapter is interested in the intersection of jargon and polysemy. In particular, the question is: how can we identify whether the meaning associated with a particular phonological or morphological string is unique to a specific specialist subgroup when that same phonological or morphological string also appears in non-specialized settings? Examining this question serves two purposes. First, discussion here will clarify discussion of “intuition” in the next chapter. Second, discussion here helps forward the non-philosophical study of jargon.

Non-philosophical research into jargon has struggled to simultaneously (a) delineate between jargon and non-jargon while also (b) having transparent conditions for doing so. In many cases, studies have left the interpretation of something being jargon or not jargon up to the participant (Subramaniam et al., 2017) or left it up to the experimenters (Bullock et al., 2019) without giving adequate explanation for how words were categorized. On the other hand, studies that have tried to categorize jargon transparently rely on a basic but flawed assumption about jargon: what makes a word or phrase jargon depends on the frequency of a *lexeme* (that is, the word or phrase – including its conjugations) in everyday English.

6.2.1 Jargon is More than Lexeme Frequency

To discuss the problem with the approach of identifying jargon with a particular lexeme, I consider the De-Jargonizer, an online tool (<http://scienceandpublic.com>) developed by a science-communication team (Rakedzon, Segev, Chapnik, Yosef, & Baram-Tsabari, 2017) meant to identify jargon in passages of text.

The De-Jargonizer takes user-entered strings of text and color codes words based on their degree of jargon-ness. Black is fine, orange is potentially problematic for low-literacy readers, and red is jargon (see Figure 6.1). The De-Jargonizer categorizes words by color based on their frequency in a corpus built from a quarter of a million articles from the BBC's website. Taking this data, Rakedzon et al. (2017) sorted word-types (where "cat" and "cats" are the same word-type) into high, medium, and low frequency based on the rate of their appearance in BBC articles. The 2000 most common word-types in corpus data, which appeared over 1000 times among the quarter million articles, were designated high-frequency words. This vocabulary size corresponds roughly to what a low-literacy speaker can be expected to know (Hu & Nation, 2000). Mid-frequency words appeared between 50 and 1000 times in the corpus. This corresponds to the 2000 to 9000 most frequent word types, which represents the level of literacy needed to understand 98% of the words in classic literature such as *The Great Gatsby* or *Lady Chatterley's Lover* (Nation, 2006). Low-frequency words, which the authors of the De-Jargonizer label as jargon, are any word type that appeared less than 50 times in the corpus of 250,000 BBC articles.

The lessons from the previous chapters' discussion of polysemy are important here. English words are not used in a single way by a single community. Instead, word uses branch off of one another, and sometimes the branches are unique to a particular community. What makes jargon jargon or technical language technical, as opposed to "folk" or "standard" language, is not the specialization of the lexeme. Instead, jargon-ness depends

The possibility of experience is, then, that which gives objective reality to all our a **priori cognitions**. Now experience depends upon the **synthetical** unity of **phenomena**, that is, upon a **synthesis** according to **conceptions** of the object of **phenomena** in general, a **synthesis** without which experience never could become knowledge, but would be merely a **rhapsody** of **perceptions**, never fitting together into any connected text, according to rules of a thoroughly united (possible) **consciousness**, and therefore never subjected to the **transcendental** and necessary unity of **apperception**. Experience has therefore for a foundation, a **priori** principles of its form, that is to say, general rules of unity in the **synthesis** of **phenomena**, the objective reality of which rules, as necessary conditions even of the possibility of experience can which rules, as necessary conditions—even of the possibility of experience—can always be shown in experience. But apart from this relation, a **priori synthetical propositions** are absolutely impossible, because they have no third term, that is, no pure object, in which the **synthetical** unity can **exhibit** the objective reality of its **conceptions**.

Figure 6.1: De-Jargonizer output for Kant’s *Critique of Pure Reason*

also on the strangeness of the way the lexeme is used. Because of this, it is a mistake to reduce, as Rakedzon et al. (2017) do, to word frequency.

To see this, consider how categorizing a word by its frequency in standard language is neither necessary nor sufficient for something being jargon. When given a passage from *Kant’s Critique of Pure Reason*, the results are unsurprising (Figure 6.1). The algorithm has picked out words like “apperception”, “synthetical”, “transcendental”, “conceptions”, and “proposition”. These are words that many philosophers would struggle to define, to say nothing of defining them in the context of Kant. The algorithm has over-generated by picking out “rhapsody”, a word that is merely old-fashioned. Not too much should be read into this specific overgeneration, however, because the mistake is caused by using a contemporary corpus to judge the frequency of words in a historical work. If Rakedzon et al. (2017) had used a historical corpus from the year of this translation of Kant, the De-Jargonizer would presumably not flag “rhapsody” as jargon.²

But we can start seeing the frequency approach’s deeper weaknesses with strings like “Pontificating about rarefied linguistic phenomena at my typographical workstation, I surmise the indubitable utility of jargon” (Figure 6.2). This sentence is merely a complicated way of saying “Thinking at my laptop, I decide jargon is useful.” There is nothing uniquely technical here, although writing and talking this way is likely to confuse and alienate many English speakers. The BBC writes plainly, and avoids pompous and obtuse

² This observation is backed up by Google Ngram’s historical tracking of “rhapsody”. See *Google Books Ngram Viewer of “Rhapsody”* (2020).



Pontificating about rarefied linguistic phenomena at my typographical workstation, I surmise the indubitable utility of jargon

Figure 6.2: De-Jargonizer output for rarefied words

language. Therefore, pompous and obtuse words are picked out by the De-Jargonizer as jargon despite failing to fulfill the De-Jargonizer’s chosen definition of jargon from the Merriam-Webster dictionary of “the technical terminology or characteristic idiom of a special activity or group” (Rakedzon et al., 2017, 2).

Not all methods of corpus-based jargon identification misidentify pompous vocabulary as jargon. Sharon and Baram-Tsabari (2013) identify jargon by comparing frequencies across multiple corpora. To simplify matters slightly, Sharon and Baram-Tsabari compare word-frequencies in a corpus of scientific discussion against a corpus of British English containing some formal settings such as government meetings. Both corpora contain speech by highly educated people, letting the authors filter out merely fancy language from jargon. Comparing the two corpora, words that appear as frequently in the stilted non-scientific corpus as the scientific corpus were identified as normal rarefied words. Words appearing more frequently in the scientific corpus were then identified as jargon. Sharon and Baram-Tsabari (2013) therefore avoid the issues with the De-Jargonizer shown in Figure 6.2.

The comparative method is not perfect, and by their own admission, Sharon and Baram-Tsabari (2013) fail to adequately categorize jargon built out of standard English words, such as “string theory” and “the big bang”. This is an issue shared with the De-Jargonizer. Consider Figure 6.3, the De-Jargonizer output for “While valid, the standard form argument begged the question”. The De-Jargonizer algorithm finds that all the words except “begged” are high-frequency words. Even “begged” is a mid-frequency word and therefore not in the category Rakedzon et al. (2017) equate with jargon. The sentence, however, is carefully constructed to contain three distinct semantic units, *valid*, *standard*

While valid, the standard form argument **begged** the question

Figure 6.3: De-Jargonizer output for an innocent and completely understandable sentence *form argument*, and *begs the question*. All three terms would not be understandable to someone without some training in philosophy or logic. Each term offers a different lesson, so each will be examined in turn.

6.2.2 Three More Lessons about Jargon

“Valid” is exactly the sort of polysemy discussed in the previous chapter. “Valid” has a number of – erm – valid uses in every day English that differ significantly from the way philosophy students are taught to use “valid” in the philosophy room. The Collins English Dictionary gives three definitions for “valid” (*Valid Definition and Meaning*, 2020):³

- A valid argument, comment, or idea is based on sensible reasoning (*They put forward many valid points for not exporting*).
- Something that is valid is important or serious enough to make it worth saying (*Most designers share the unspoken belief that fashion is a valid form of visual art*).
- If a ticket or other document is valid, it can be used and will be accepted by people in authority (*For foreign holidays you will need a valid passport*).

“Valid” is polysemous in these three ways among the general public, but it has a fourth polysemous use in the context of philosophy. Roughly, to be valid is to have a logical form where the conclusion follows from the premises. There are many competing ways to

³ The choice of Collins over the OED, despite my reliance on the OED in the previous chapter, is deliberate. Collins is a contemporary, or synchronic, dictionary whereas the OED is historical dictionary. Given I am discussing contemporary uses of “valid”, Collins is therefore preferable for present purposes.

clarify this, but nonetheless, this philosophical use is related but distinct from the uses described by Collins.

The problem caused by the polysemy of “valid” is well known to anyone who teaches first-year undergraduates or non-philosophers. Student essays often contain jarring uses of “valid” such as “the argument was sound but not valid” or “while valid, the conclusion does not follow from the premises”. Such sentences show ignorance, but it is important to note what exactly that ignorance is. It is not a mistake about the nature of validity in the same way that the metalogic theory *an argument is valid if it is true in all models of logic* would be a mistake if false. Rather, undergraduates hear their teachers use “valid” but fail to recognize that their teachers mean something different than everyday English. Connecting this discussion of “validity” with the previous chapter, we get the following:

SECOND LESSON ABOUT JARGON: Otherwise everyday English words can have polysemous meanings that are jargon.

While “valid” is a straightforward extension of last chapter’s discussion of polysemy, “standard form argument” is not. There is nothing relevantly polysemous about any of the words, nor does the compound phrase have a meaning outside of philosophical settings.⁴ The phrase seems to act like a normal descriptive phrase, where the meaning of the phrase is a product of the meaning of the words. A standard form argument is an argument that has been converted to a form that is standard. But even understanding the constituent words and understanding how they combine is not sufficient to know the phrase’s extension, intension, or its effect on a sentence’s truth conditions.

Linguists have a name for the relationship between the meaning of parts and wholes, SEMANTIC TRANSPARENCY (or sometimes just TRANSPARENCY), and its opposite, SEMANTIC OPACITY (Libben, Gibson, Yoon, & Sandra, 2003; Plag, 2003). Transparency

⁴ Using both my own imagination and searches for “standard form argument” on Google, I cannot find a way to use the string “standard form argument” except to talk about standard form arguments. This is certainly not to be taken as proof that no such sentences can or do exist.

and opacity are the degree of relationship between the meaning of a compound noun or phrase and the meaning of its parts. More precisely, Schäfer (2018) defines transparency as

A measure of the degree to which the meaning of a multimorphemic combination can be synchronically related to the meaning of its constituents and the typical way of combining the constituent meanings (2018, 1).

Breaking that down, transparency is a property of words or phrases that can be broken down into meaningful parts. This might be compound phrases such as “the big bang”, compound nouns like “mailman”, or other words that can be broken down into multiple meaningful parts (i.e., morphemes) like “nonflammable” which has 3 parts: “non”, “flamm”, and “able”.⁵ The transparency of a word or phrase comes in degrees, running from transparent to opaque based on the typical ways constituent parts combine to form meanings. Finally, transparency only measures how connected they are in the eyes of modern language speakers and does not take into account historical uses of the words or other etymology.

To illustrate the notion, compare two compound constructions of “hot” and “dog”:

(39) It is a hot dog.

(40) It is a hotdog.

(39) contains some ambiguity, but the ambiguity is limited and due to ambiguity in one of the constituents. Is the dog popular, sexually appealing, or overheated? Otherwise, the noun-phrase “hot dog” works like any other combinatorial noun-phrase as the meaning of the noun phrase comes from the adjective modifying the noun (Millar & Trask, 2015, 27). Compare that to (40). Historically, “dog” was American slang for a sausage, and “hot dog”, later shortened to “hotdog”, was a modification of this sense of “dog”. However,

⁵ See Chapter 2 of Plag (2003) for a discussion of the theoretical difficulties behind the notion of morpheme or parts of meaning.

this meaning of “dog” has fallen out of contemporary circulation (except, curiously, in American English as a shortened form of “hotdog”). So to the modern English speaker, especially those outside of America, there is almost no apparent relationship between the meaning of parts “hot” and “dog” and what hotdogs are.

“Standard form argument”, despite looking like a descriptive phrase, is opaque. In fact, it may only apparently be a descriptive phrase. Instead, it might be a calcification of a demonstrative reference-fixing phrase. Plausibly, when first used, “standard form argument” picked out the form of arguments that was standard in the context in which the phrase was coined.⁶ The phrase stuck as the name for a certain type of argument even when it is not the standard form of argument in that context. This then leads to yet another conclusion about jargon:

THIRD LESSON ABOUT JARGON: Not only does jargon include specialized polysemous meanings of words, but jargon also includes specialized combinations of standard English words.⁷

The final jargon-element from above, “begged the question,” is an interesting example in light of transparency because “begged the question” is only apparently transparent. The non-philosophical use of “begs/begging/begged the question” is newer than the philosophical use, which dates to a 16th century translation of Aristotle (*Beg the Question*, 2016).⁸ In contemporary usage, begging is roughly equivalent to asking, although begging has connotations having to do with the dignity of the act. Therefore, the combination of the three elements, when combined in the normal way, would mean, roughly “to ask

⁶ In theory, I am against this sort of armchair historical lexicography, but the example is only illustrative. I looked into the etymology of “standard form” to described numbered arguments, and I cannot find a use prior to the early 2000s.

⁷ In discussion of jargon, Sharon and Baram-Tsabari (2013) makes a similar observation in passing about terms like “the big bang” without reference to reference-fixing descriptions or linguistic transparency and opacity.

⁸ The most recent textual example of the philosophical sense of “beg the question” recorded by the OED2 is delightfully stuffy: From “F. C. Bowen *Logic* ix. 294 (1870): ‘The vulgar equivalent for *petitio principii* is begging the question.’ ”

for the question”. While it is not immediately obvious how one person could ask for a question, given what rhetorical moves might involve asking for questions, it is entirely reasonable that people hearing the phrase without knowing what it means began to infer the phrase denotes implicitly inviting a question. This interpretation is more closely related to typical way the multimorphic phrase “begging the question” would combine than the way philosophers and lawyers use the phrase to talk about the fallacy of assuming an argument’s conclusion in the argument’s premises.

Here we find an important mechanism for confusion and miscommunication relating to jargon, as it relates to people hearing the word and guessing what it means:

FOURTH LESSON ABOUT JARGON: Jargon can be appear to mean one thing while actually meaning another

The four lessons are important for what follows. Jargon does not just exist as a barrier of communication between members of disciplines and people who are not members of the discipline. Jargon can be localized to many different types of social groups within disciplines, such as intellectual traditions, specific debates, or subfields (Lesson 1). Jargon is not just limited to fancy-sounding words like “polysemous”, “doxastic”, or “*prima facie*” (Lesson 2). Jargon may be driven either by special polysemous meanings of a word or combinations of everyday words whose meaning is not what the words would standardly mean when combined in the way they are combined (Lesson 3). Sometimes the lack of apparent fanciness leads to confusion. Because jargon can look like it means one thing while having another meaning only known to experts (Lesson 4), certain technical language can be apt for miscommunication.

6.3 Polysemy Jargon

The connection between the previous chapter’s discussion of polysemy and this chapter’s discussion of jargon should now be apparent. On one hand there is jargon that wears its inaccessibility on its sleeves by using strange or obviously opaque lexical items. Call this jargon OPAQUE JARGON. On the other hand, there is jargon that looks like common language because it uses common words in specialized but related ways. Call this POLYSEMY JARGON.⁹

On semantic externalist accounts of language, besides the semantic and pragmatic information carried by words, many words have what Frege called COLORING (Frege, 1997; Sander, 2019) or what Cappelen has recently called LEXICAL EFFECTS (2018, Ch 11). These are mental images, associations, behavior, etc., caused by a lexical item that are not reducible to what is said or communicated by the lexical item. As an example, imagine I tell a friend “I am getting coffee this afternoon with Natalia”, unaware that my friend does not know who Natalia is. My friend may reasonably infer I am getting coffee with a woman. They also even reasonably infer that the woman is of Eastern European heritage. This is not contained in semantic information or nor is my utterance communicating anything about Natalia’s heritage. Instead the lexical item “Natalia” carries certain information driven by induction. Because people we encounter called “Natalia” tend to be women and often have Eastern European heritage, my friend makes a well-founded inference about the referent of the name.

Normally, opaque in-group language such as opaque jargon or opaque slang has the lexical effect of conveying group-membership or exclusion from a group. Using fancy technical language intentionally or unintentionally signals to non-expert listeners that they lack the prerequisite knowledge, vocabulary, or other form of expertise in order to understand

⁹ This distinction ignores the possibility of jargon that uses a common phonological or morphological string but in an unrelated way. Arguably “begging the question” falls into this category. I am not interested in this category at the moment, but much of what is said about polysemy jargon applies because it can still be checked for using the tests outlined for below.

what is said.¹⁰ This is not the case with polysemy jargon, and one lexical effect important to the present discussion of polysemy jargon is what we might call JOINT-SAYING. When two people use the same lexical item in speech, the default assumption by members of the conversation is that they mean the same thing by the lexical item. We expect people in our linguistic communities to speak the same way we do. Not having the defeasible but default assumption that people mean the same thing when they use the same words would make communication nigh impossible as conversational participants would need to rule out – as much is possible – interlocutors having fundamentally different idiolects.

Polysemy jargon can cause both conversational participants to believe in joint-saying when they do not mean the same thing by the word. This is why reading a philosophy student say “the argument is sound but not valid” is so jarring. As people used to using “valid” in a specific way unique to logic and philosophy, we expect “valid” in philosophy contexts to mean something related to logical entailment. Given we have used the polysemy jargon meaning of “valid” in philosophical contexts with students, we expect they will use it in that way in return, and because many students do not know there is a specialized meaning of “valid” they assume we mean it in a normal way. We both infer joint-saying about “valid”, despite meaning different things by it.

In the context of philosophy, it is reasonable for us to assume a speaker means *the conclusion is logically entailed by the premises* by “valid” because that is what philosophers typically mean by “valid”. I am not suggesting though that “valid” is contextual in the way indexicals like “I”, “she”, and plausibly “tall” are (D. Kaplan, 1989).¹¹ There are plenty of valid uses of “valid” in philosophical contexts that do not mean (roughly) logical entailment. Instead, because the polysemy of words like “valid” forces us as listeners to disambiguate between multiple meanings, in various contexts we expect the meaning most common in those contexts unless we have reason to think otherwise. In this way,

¹⁰ I do not mean for this to be a particularly robust inference that listeners make. The listener either spends effort attempting to interpret the utterance or they do not.

¹¹ In other words, I am not a radical contextualist like Searle (1978), Recanati (2004), or others.

context plays a role in disambiguation but does not affect the meaning of the word.

Because polysemy jargon is both limited to certain social groups and, unlike opaque jargon, apt to falsely cause the lexical effect of “joint-saying”, it is apt to cause verbal disputes. We should especially be worried about this having an effect on philosophical discourse. Philosophy is full of technical discussions with common English words like “knowledge”, “good”, “right”, “belief”, and “concept” at the center of discussion, meaning there is ample opportunity for polysemy jargon to subtly develop. For this reason, for metaphilosophical purposes it is important to develop tools to test for polysemy jargon, and the existing corpus-based methods described above in Section 6.2 cannot do it. While opaque jargon will be flagged by corpus-based methods like the De-Jargonizer because opaque jargon has a unique lexical item, polysemy jargon will not be flagged. It is time, then, to look at methods that can spot polysemy jargon.

6.3.1 Anaphors and Zeugmas

Standard armchair tests for polysemy can spot polysemy jargon, but only if they are evaluated by people with some acquaintance with the polysemy jargon, such as experts. Linguists have developed tests to determine whether or not a potentially ambiguous word has two meanings that rely on awkward-feeling uses of zeugmas and anaphors (Vicente & Falkum, 2017). A zeugma is a rhetorical tool where a word is used and then suppressed later in the sentence. Consider:

(41) We were both in the woodworking class. The chair I made was plain and awkward,
and her chair was ornate and clean.

(42) We were both in the woodworking class. The chair I made was plain and awkward,
hers ornate and clean.

In (42), “chair” is suppressed when discussing the second person’s chair, but the sentence feels felicitous and clear. Consider instead:

- (43) We both took our own morning walks past our favorite bank – hers pebble-strewn and muddy, mine brutalist and indicative of the great wealth kept inside.

Because “bank” is used in a way that conflates two distinct meanings, the appositive strikes us as wrong and awkward. The zeugma test is then taken to indicate that “bank” is ambiguous between two meanings. The zeugma test therefore uses an evaluator’s linguistic competence to check whether a suppressed use of a word feels awkward when modulating between two purported senses.

Consider now a zeugma involving the unambiguous polysemy jargon example of “string”, to determine if physicist’s use of “string” has a different meaning than non-specialist use of “string”:¹²

- (44) I am a quantum physicist, and I am married to a knitter. He spends time working with strings, as do I.

The swap feels awkward and strange. Thus we can take it as evidence that there are two meanings of string.

A similar test involves using anaphoric reference to an earlier word:

- (45) Unlike my awkward and plain chair, she made one that was ornate and clean.
(46) She passed a bank teeming with wildlife, and I passed one full of tellers.

In both cases “one” gains its meaning from the previous clause. In (46), a similar awkwardness arises due to “bank” being used in two different ways. The anaphor test produces similar awkwardness with polysemy jargon involving “string”:

- (47) Turns out that the newlyweds both work with the same thing! The husband, a knitter, uses strings, the wife, a quantum physicist, studies them.

¹² The example of “string” is from Plunkett and McPherson (2019).

While sometimes criticized (see Geeraerts, 1993), anaphors and zeugmas are generally-accepted tests for polysemy (Vicente & Falkum, 2017), and the anaphor and zeugma test of “string” both result in strange-sounding and strange-feeling sentences.

In the case of polysemy jargon, only specialists or people with some competence in specialist language are qualified to judge anaphor and zeugma tests because both tests require competence with the different polysemous meanings. Imagine that you do not know of the use of “string” to refer to the theoretical entity posited by string theorists. Consider (44) and (47) again. The sentences still feel strange, but in a different way. Rather than not quite being understandable or felicitous-feeling, their truth conditions are strange. After all, why would a quantum physicist study yarn? Without knowing the jargon meaning of “string”, the sentences lack plausible interpretations, but it may gain a plausible interpretation if we learn new things about the world, such as quantum physicists’ discovery of strange phenomena involving wound string.

Full expertise is not necessary to run the anaphor or zeugma tests, however. My entire knowledge of strings is contained in the sentence, “They are very small and very thin fundamental entities posited by string theory made up of pure energy, and some physicists are not sure their existence can even be tested”. What seems to be required is merely the knowledge that the word has a specialized sense and some rough idea of what that specialized sense refers to. This latter criteria seems to be close to what Carey calls *PLACEHOLDER CONCEPTS* (2011, Ch 8-13). Placeholder concepts develop when a word is used in a way that does not make sense or people are exposed to phenomena inexplicable to their current conceptual structures.¹³ Complex concepts can take anywhere from hours to years to develop, and placeholder concepts form in the intermediary. Placeholder concepts represent the relationship between the concept and other concepts without full representation (or in philosopher-ese, understanding) of the thing the concept represents.

¹³ Being a developmental scientist, Carey is using “concept” to mean token mental representations as opposed to the abstracta that those representations represent. Anyone who does not like this usage of “concept” is free to substitute “concept” for “mental representation” or “conceptual competence”.

So, while I know that strings have some relationship to string theory, energy, and quantum phenomena, I lack a proper representation of strings that represents the relationship between it and other entities posited by physicists. A quick perusal of scholarly articles in string theory suggests I do not understand the majority of sentences involving the jargon sense of “string”, but because I know the term is being used in a special way related to fundamental physics, I know not to try to interpret sentences as being about macro-level fabric. This, it seems, is enough competence to be able to spot polysemy in the anaphor and zeugma tests above.

The need for partial competence to check for polysemy-jargon comes with downsides, however. This test is not much help if we are not specialists and we are not sure if there is polysemy jargon in the first place. This can be exacerbated when the specialist/non-specialist distinction does not fall under expected social lines, as we may not know we are not a specialist in the relevant way.

To see this first-hand, consider the following sentence inspired by anthropology, which is about Kula objects – objects used in market-like exchange systems in the mid-Pacific Massim Archipelago:¹⁴

- (48) Kula objects are a material-based reflection of epistemic traditions surrounding economy and social hierarchy.

In the analytic philosophy meaning of “epistemic”, the sentence is difficult to understand. Analytic philosophers use “epistemic” in a way that does not readily admit to traditions, at least not traditions that exist outside of scholarly traditions. Moreover, these traditions are certainly not the sort of thing that readily admit to material-based reflections, especially not objects used in exchange systems.

The truth-conditions of the sentence seem strange to many analytic philosophers because “epistemic” has both the sense analytic philosophers are used to and a sense more

¹⁴ Thanks to Sonja Dobroski’s help coming up with this example.

common among critical theorists who draw upon Foucault. “Epistemic” in (48) corresponds roughly to the way Kuhn (1996) uses “paradigmatic” or to Quine’s discussion of conceptual systems if extra weight is added to the role of history of those conceptual structures (Cooper, 2007, 15). People not aware of this non-analytic meaning of “epistemic”, however, are apt to misinterpret the sentence as meaning something it does not.

Hopefully (48) helps drive home the point I am making about miscommunication led by polysemy jargon. As experts we have mastery over certain words and their meaning. Nonetheless, these words sometimes have meaning we do not know about. If we are not on the lookout for polysemy jargon and our interpretation of the sentence makes no sense or seems to have outlandish truth conditions, we may assume joint-saying and accuse the speaker of saying something false, ignorant, or confused. This is the main challenge of polysemy jargon. Unless we have reason to suspect it is there, use of an unknown sense may look like false metalinguistic or object-level beliefs.

In light of the zeugma and anaphor test requiring partial competence with the jargon polysemy, I will now suggest three methods that do not require partial competence with expert vocabulary in order to spot potential polysemy jargon. I briefly highlight methods in corpus linguistics other than the methods described above, argue that philosophers’ use of “sense” is indicative of jargon polysemy, and advocate that in order to spot polysemy jargon, ultimately philosophers need to do the hard work of lexicography themselves. Like the zeugma and anaphor test, each test is not meant to be ultimate arbiters for whether or not a term is low-key polysemous in the mouths of specialists. Rather, I suggest tests that can be refined moving forward, and, in the case of “sense”, demonstrate that the methods I have used to spot jargon polysemy do not overgenerate to all terms used in specialized circumstances.

6.3.2 Corpus Linguistics

Corpus linguistics has gained popularity among experimental philosophers in recent years (Hansen et al., 2019; Mizrahi, in press; Nichols & Pinillos, 2018; Ramos et al., 2019). These studies use corpora of natural language analyzed en masse. These studies either analyze the data using armies of human coders or computer programs that identify statistical patterns in how certain words are used. Many corpus linguistics studies are in some sense lexicography done at a large scale; instead of a handful of lexicographers finding definitions for a handful of representative uses of a word, corpus analyses look at how words are used in thousands of occurrences in natural language.

Most methods of corpus linguistics are not fit for task here. As discussed above, many existing methods employed in the hunt for jargon only focused on lexeme frequency, which cannot spot polysemy jargon. Instead, identifying polysemy jargon requires corpus linguistics methods developed to sort out meanings. This can be hand-coded based on prior linguistic competence of human analyzers, as was done with different uses of “know” by Nichols and Pinillos (2018). This requires researchers to have a prior understanding of what the different senses or uses could be. This gives hand-coding the same problems as the anaphor and zeugma test, since it requires prior competence with the polysemy jargon.

Alternatively, there are computational methods, specifically vector space analysis, as used by Fischer, Engelhardt, and Herbelot (2015). The specifics of the theory are well beyond my limited knowledge in mathematics and coding, but the gist of vector space analysis is that it infers similar meaning from co-occurrence of words. Words like “laptop” and “keyboard” will co-occur, signaling closeness of meaning, but “keyboard” will appear more frequently than “laptop” in relationship to “letters” or “typing”, suggesting “keyboard” is more closely related in meaning to those words than “laptop”.¹⁵ As described,

¹⁵ Those interested can try their hand at reading Turney and Pantel (2010) for more information. See also Kosem (2016) for a more accessible introduction to computational methods that still require human

vector space analysis assumes univocal meanings of words, but Fischer et al. (2015) uses vector space analysis to determine which polysemous sense of “appear” and “seem” is most prevalent in natural language. There may then be ways to expand the method to help identify whether a term has an unknown specialized polysemous sense.

Besides methodological hurdles, corpus linguistics is also either technically demanding or time consuming and expensive. Corpus linguistics either requires a large budget to pay for human coders or significant technical expertise. In cases where a suitable corpus is not already built (e.g., JSTOR or one of the many natural language corpora built specifically for corpus analyses), even getting to the analysis stage can be time-consuming. This can involve scanning the required documents, converting the scans into text, then, if the scan-to-text conversions uses software, proofing and correcting the output by hand. For this reason, while corpus linguistics has been employed in the use of polysemy in philosophy (Fischer et al., 2015), barriers stand between it and its use to spot polysemy jargon.

6.3.3 Sense-Talk

While much attention has been given to intuition-talk, both in this dissertation and elsewhere (see Section 7.2 below), the patterns behind intuition-talk are fairly common. There are other words philosophers love to use for certain important argumentative functions that are used without those philosophers having clear definitions or accounts in mind. In this section, I will explore whether one of those unreflectively used words, “sense”, is indicative of polysemy (and thus jargon polysemy) when used by philosophers. Here, I am not interested in all uses of “sense” by philosophers, since “sense” is often used in relationship to sense-experience and as a translation of Frege’s *Sinn* (1948). Instead, I am interested in “sense” when it is used as a tool of disambiguation, where it used in such ways as “in the author’s sense of P...” or “in the normative sense of Q...”.

analysis.

Consider the two top search results on the Stanford Encyclopedia of Philosophy (plato.stanford.edu) for “sense” after removing hits discussing sense-data and other perception-like processes.¹⁶ From the “Representational Theories of Consciousness” entry [emphasis added]:

The notions of consciousness most commonly addressed by philosophers are the following: (1) Conscious awareness of one’s own mental states, and “conscious states” in the particular *sense* of: states whose subjects are aware of being in them. (2) Introspection and one’s privileged access to the internal character of one’s experience itself. (3) Being in a sensory state that has a distinctive qualitative property, such as the color one experiences in having a visual experience, or the timbre of a heard sound. (4) The phenomenal matter of “what it’s like” for the subject to be in a particular mental state, especially what it is like for that subject to experience a particular qualitative property as in (3). Block (1995) and others have used “phenomenal consciousness” for *sense* (4), without distinguishing it from *sense* (3). (Lycan, 2019)

Lycan seems to be using “notion” and “sense” interchangeably here to distinguish these four notions/senses of consciousness. These, as the author explains at the end, can be discussed in at least partial isolation from each other and are sometimes confused. Moreover, according to the author one sense has a name (“phenomenal conscious”) that does not pick out what is discussed in at least two of the three other senses. So “sense” in the passage is being used to disambiguate things that share the name “consciousness”, are potentially metaphysically distinct, but are also nonetheless related in subject-matter.

¹⁶ This is meant to be a pseudo-random process. The SEP was chosen over other repositories like PhilPapers because it allows for easy in-text search and is not behind a paywall. I removed one top hit (“Mysticism”) in favor of “Representational Theories of Consciousness” for ease of presentation because in the article on mysticism the use of “sense” considered in this section appeared in the same passage as a quasi-perceptual use of “sense”.

“Sense” is used almost identically in the introduction of “The Definition of Morality” entry of the Stanford Encyclopedia of Philosophy [emphasis added]:

There does not seem to be much reason to think that a single definition of morality will be applicable to all moral discussions. One reason for this is that “morality” seems to be used in two distinct broad *senses*: a descriptive *sense* and a normative *sense*. More particularly, the term “morality” can be used either

1. descriptively to refer to certain codes of conduct put forward by a society or a group (such as a religion), or accepted by an individual for her own behavior, or
2. normatively to refer to a code of conduct that, given specified conditions, would be put forward by all rational persons.

Which of these two *senses* of “morality” a theorist is using plays a crucial, although sometimes unacknowledged, role in the development of an ethical theory. If one uses “morality” in its descriptive *sense*, and therefore uses it to refer to codes of conduct actually put forward by distinct groups or societies, one will almost certainly deny that there is a universal morality that applies to all human beings. (Gert & Gert, 2017)

Similar to the “Representational Theories of Consciousness” entry, “sense” is disambiguating two things that may be metaphysically distinct but are related in subject-matter and by the word “morality”. Notice too that this passage explicitly endorses at least a use-polysemous account of “morality”. According to Gert and Gert (2017), morality can be used in two consistent, distinct, and related ways. Given that these are different ways of using “morality”, theories about either sense differ accordingly. The authors therefore seem to actually be committed to at least belief-polysemy or implicature-polysemy since

the authors think that some sort of mental state is driving these patterns. However, before the quoted passage, the authors explicitly deny that the concept/conception distinction is relevant here, which would have otherwise allowed the authors to say that the definitions were different conceptions (that is, beliefs) about the same concept. Therefore, it looks like Gert and Gert (2017) are committed to meaning-polysemy of “morality” and are using “sense” to disambiguate polysemous meanings.

The most straightforward way of reading the use of “sense” in the two passages are that the philosophers writing the articles are using “sense” in the standard English way of referring to distinct meanings of a word. Call this the **STRONG VIEW OF “SENSE”**:

“Sense” disambiguates different (literal) meanings of a word.

If the Strong View of “sense” is true, then sense-talk is a guide to polysemy, and therefore polysemy jargon. Since we need some way of signaling modulations between different meanings of a word when the ambiguity may cause confusion, according to the Strong View of “sense”, “sense” allows us to do this. Therefore, anytime we see sense-talk used as a disambiguation tool, we have strong evidence of polysemy. Going a step further, if philosophers are distinguishing between senses that are not found in standard English, then we have evidence of polysemy jargon. This does not make spotting polysemy jargon trivially easy. False semantic beliefs may lead to inappropriate sense-talk, and determining what senses exist in standard English is not as simple as checking the dictionary. Dictionary definitions of philosophically relevant terms often have philosophical assumptions baked into the entries, and no dictionary is an infallible guide to English.

There is weaker way of interpreting philosopher’s sense-talk. Perhaps I read too much into the apparent metaphysical differences between different senses above, and in fact there is only one correct or true meaning of “consciousness” and “morality”. “Consciousness” refers to something in philosophical settings, and it is consciousness. “Morality” refers morality. These so-called “senses” are just different theories about the subject, and

at most one is right.

Call this objection the WEAK VIEW OF “SENSE”:

“Sense” signals modulation between belief-level phenomena such as different conceptions or beliefs about a word’s meaning.

It should be clear now why the Weak View of “sense” is wrong. It ignores the fundamental polysemous nature of language. As I defended at length in Chapter 5, words sometimes mean one thing when spoken by one person at one time and another thing when spoken by that person or another person at another time. In normal English contexts, sense-talk helps us communicate when we are changing between one meaning of a word and another (see Atkins & Rundell, 2008, 270). Unless we have reason to suspect “sense” has a jargon-polysemous sense, we should therefore take sense-talk at face value.

The zeugma and anaphor tests seem to indicate that “sense” is not polysemy jargon in the mouths of philosophers. Here is a zeugma and an anaphor where one use of “sense” is meant to be in the way Lycan (2019) and Gert and Gert (2017) use it (remembering they are taken to be representative of philosophical uses of “sense”) and the normal English use of “sense” to mean a meaning. In the following zeugma, the philosopher use of “sense” comes first, followed by a suppressed standard English use of “sense”:

(49) Gert and Gert (2017) describe two senses of “morality”, the Collins English Dictionary three.

In the following anaphor, the standard English meaning of “sense” comes first, followed by the philosopher’s use:

(50) “Consciousness” is used in different ways. The OED3 gives one sense of the word as *The state of being aware of and responsive to one’s surroundings*. Lycan (2019) suggests one relevant to philosophy is *Introspection and one’s privileged access to the internal character of one’s experience itself*.

Neither the zeugma or anaphor have the “off” feeling of the other zeugma and anaphor tests used above. Instead (49) and (50) seem to be perfectly adequate uses of anaphors and zeugmas, and the tests give no reason to think “sense” here means anything different in philosophical contexts than it does in standard English. Therefore, just like in non-technical settings, disambiguating sense-talk is a guide to polysemy jargon. Moreover, it is a method to finding polysemy jargon that non-experts can use without technical barriers.

6.3.4 The Semantic Anarchy Worry

It is worth briefly pausing to highlight the importance of the zeugma and anaphor tests supporting the Strong View of “sense”. A potential worry about the framework developed here is that it entails semantic anarchy, where specialist language is unmoored from everyday language.¹⁷ According to this worry, I develop a picture where philosophers speak a language that merely looks and sounds like standard English but has little to no semantic overlap among theoretically important or theoretically contentious words.

This is not the case. I am instead defending the view that philosophers *sometimes* use theoretically important words with a different meaning than non-philosophers. Philosophers sometimes identify important categories, kinds, or distinctions that are beyond the expressive power of everyday English. To refer to such things, philosophers often knowingly or unknowingly co-opt existing words, and if the linguistic innovation takes off, philosophers thereby use words with a polysemous meaning whose circulation is limited primarily to other philosophers. How often that happens is unclear and something I am not prepared to try to estimate. I do believe, however, that any time philosophers in a debate cannot agree on even what sort of metaphysical category a thing is – such as the disagreement over whether intuitions are an epistemic kind or a cognitive kind or

¹⁷ Thanks to Brian Weatherson for first putting this worry in my head in discussion of a *very* early version of the idea presented here.

the disagreement over whether concepts are token psychological entities or metaphysical abstracta – is likely a place polysemy exists or jargon polysemy has developed.

I return to this worry in more detail in Section 7.4, but for now I want to highlight how the zeugma and anaphor test for “sense” pushes against the worry about semantic anarchy. Philosophers use “sense” in a technical setting to help disambiguate *something*. It might be something philosophy-specific such as between conceptions or different philosophical accounts. The tools I develop here indicate it disambiguates the same thing “sense” does in standard English – polysemous meanings of a word. As discussion of “sense” demonstrates, I am not committed to the claim all technical-seeming words used by technicians in technical settings is polysemy jargon. Whether a word is polysemy jargon needs to be taken at a case-by-case basis, and my account correctly handles this.

6.3.5 Lexicography

The final method, and the one to be employed in the next chapter, is to apply the methods of lexicography to texts in philosophy. The methods of lexicography, while becoming more automated (see Kosem (2016)) still require careful manual examination of a word’s use and consideration of the context surrounding its use (Atkins & Rundell, 2008; Knowles, 2010). Anyone looking for an exact algorithm or even concrete guidelines on how to convert textual evidence to senses or dictionary definitions will be disappointed. While it is easy to find guides on what sort of evidence one should compile and consider when looking for distinct senses of a word, comparatively little exists on how to convert textual evidence to senses. Thinking among lexicographers seems to be that lexicographers rely largely on their own linguistic competence. Just as astronomers may see telescopes as the instruments of astronomy, lexicographers see humans as one of the instruments of lexicography.¹⁸ Largely dismissive of using linguistic theory in the process of forming

¹⁸ This idea is repurposed from Weinberg (2015)

definitions, the authors of one handbook say “our experience as editorial managers suggests that good lexicographers operate to a large extent on the basis of instinct, sound judgment, and accumulated expertise.” (Atkins & Rundell, 2008, 130)

There is more to contemporary lexicography than just examining sentences and words, and modern lexicographers are also responsible for building and cleaning corpora, editorial decisions about which senses belong in a dictionary, writing those dictionary definitions, and finding representative uses of a given definition for inclusion in a dictionary. Here I am only interested in trying to tease apart what specific philosophers have meant (or tried to mean) by certain technical language. For that, I am going to set aside building corpora and writing dictionary definitions and focus on cataloging senses. For this, the same lexicography handbook quote above recommends drawing upon three sources of knowledge (where a headword is the word or phrase that a dictionary entry is organized around):

- (a) what we know, as native speakers, about the headword (its inherent properties)
- (b) what we learn from its use in corpora and elsewhere (its contextual features)
- (c) what we know about where the citations came from (the properties of the source texts) (Atkins & Rundell, 2008, 150)

In the next chapter, I consider “intuition”. As speakers of English (I am going to ignore the qualification of native speaker by Atkins and Rundell (2008)), we have a general sense of what “intuition” means. It is a reaction that involves some sort of insight and ease. Philosophers reading this know that “intuition” is used in philosophy surrounding thought experiments and typically used in support of certain claims. Philosophers reading are also competent in using “intuition” in the sort of way philosophers typically do:

(51) My intuition is that it is wrong to kill the man in this instance.

(52) My intuitions disagree with yours.

(53) I just don't have that intuition.

This competence corresponds to (a) from Atkins and Rundell (2008), namely, what we know as speakers of a language.

In terms of (b) – what we learn from a word's use in a corpora and elsewhere – I am not interested in building a corpora, nor do I have the technical expertise to do so. That does not mean I do not know anything about how “intuition” is used in philosophy. As Andow has explored (2015a; 2015b; 2017), “intuition” language is very widespread in philosophy. Moreover, anyone working in the area knows there is very little agreement over the basic facts of what intuitions are, especially among the people who have set out to pin down an account of intuitions.

Typically when lexicographers draw upon their knowledge described in (c) – what they know about where citations come from – lexicographers are checking for the regional dialect or specialized language of the author. If lexicographers observe a use of “jug” that only appears in rock climbing magazines, lexicographers have evidence that the use of “jug” is slang among rock climbers. Similarly, use limited to a single region's regional newspapers or a certain timespan suggest the sense is limited in time or space. Looking at “intuition” among philosophers requires more granularity. As I discuss in Section 7.6.1, if there are multiple related meanings of “intuition” among metaphilosophers, they are driven by theoretical necessity. Because different metaphilosophers have different theoretical goals in mind, one sense of “intuition” is not shared by all metaphilosophers. Each author has different theoretical goals, commitments, and idiolects, so each author needs to be examined individually.

6.4 Conclusion

This chapter examined the interaction of polysemy and jargon. This builds upon my analysis in Chapter 5, where I argued that polysemous words have multiple related meanings. As I discussed in this chapter, technical language often takes the form of polysemous senses. These are meanings of polysemous words only known to a small circle of experts. Words that look like they are used with a standard meaning can actually mean something quite technical. Such polysemy jargon creates a number of issues, both with the study of jargon and with communication involving jargon.

In the first part of this chapter (Sections 6.1 and 6.2), I introduced this phenomena and discussed why identifying jargon by assuming jargon only involves fancy words is a mistake. This involved critiquing existing techniques of studying jargon outside of philosophy and using them to illustrate important features of technical language. Section 6.3 introduced the notion of polysemy jargon and focused on the question of how we could identify polysemy jargon.

In this and the previous chapter, I laid the groundwork for a major part of this dissertation's positive thesis. The linguistics of philosophy is key to understanding the epistemology of philosophy. However, the linguistic study of philosophy, specifically the lexicography of philosophy, needs to be done with an appreciation for the complexities of lexicography and with correct metasemantic assumptions. As I discuss next, previous attempts to do the lexicography of "intuition" have failed on both accounts. Consequently, metaphilosophers have formed incorrect understandings of the intuition debate.

Chapter 7

Case Study: “Intuition” As Polysemy Jargon

7.1 Introduction

The first four chapters of this dissertation focused on the epistemology of philosophy, specifically discussing faults with exegesis-based intuition denial. I critiqued Deutsch and Cappelen’s intuition denial strategy by first arguing that exegesis-based intuition denial fails to avoid experimental-driven skepticism (Chapter 2). Then I argued that the method of checking original texts lacked value in debates about how thought experiments work (Chapters 3 and 4).

Chapter 5 saw a shift from the epistemology of philosophy to the philosophy of language of philosophy. The goal of Chapters 5 and 6 has been to develop a semantic externalist framework for understanding how technical language in philosophy works. Over the course of this discussion, I have slowly moved back to the issue at the heart of the dissertation, namely, the relationship between the discipline of philosophy and its practice. This arc started in Chapter 5, where I argued that established polysemous senses of a single word are distinct meanings of a word. I then combined this with discussion

of technical language in Chapter 6, drawing lessons from previous non-philosophical attempts to separate jargon from non-jargon. I argued that words, even common English words, have distinct meanings that are only known by specialists. Importantly, since these meanings work like normal polysemy, a speaker can choose to use the technical sense in any context or with any subcommunity of speakers just as I can use “mouth” to talk about the estuary down the hill from my flat *or* the inside of my cheeks.

As a case-study of the framework developed over the last two chapters, I end the dissertation by discussing metaphilosophers’ use of “intuition”. The choice of “intuition” is in part structural, since it connects the first half of the dissertation with the second. The choice of “intuition” is also partly autobiographical – spending three years trying to convince people that the intuition debate is a verbal dispute led me to develop much of the framework in the previous two chapters, and I do not want it to go to waste. The discussion of “intuition” is also partly to tie up a loose end from earlier chapters. Despite intuitions’ centrality, at no point have I taken a stance on what intuitions actually are. This chapter will argue that there is no easy or single answer to that question. Finally, the discussion of “intuition” is also to drive home the thesis of this dissertation. If metaphilosophers want to get to the bottom of how philosophy works, they need to take on a broad array of empirical data. Metaphilosophers should develop research programs drawing from sociology, anthropology, and lexicography.

To some extent, there is already a research program on the lexicography of philosophy. Philosophers, most notably Williamson (2007) and Cappelen (2012), have engaged in the lexicography of “intuition”. These discussions are hobbled by overly simplistic methods, however. They do not appreciate the complexity of language nor the complexity of lexicography. In this chapter I aim to push that program forward by building my discussion of the lexicography of “intuition” on the previous two chapters. I look at uses of “intuition” in a handful of specific metaphilosophical texts (Bengson (2015); Ludwig (2010);

Nagel (2012b); Weinberg et al. (2001)), and I argue that at least two of these texts have distinct polysemous-jargon senses of “intuition”.

In Section 7.2 I review existing attempts to look at “intuition” (Cappelen, 2012; Devitt, 2015; van Inwagen, 1997; Williamson, 2007), arguing that they fall short because they all fail to understand the metasemantic properties of jargon. Then, I turn to my own lexicographic analysis in Section 7.3, where I argue Weinberg et al. (2001) and Bengson (2015) have identifiably distinct polysemy-jargon meanings of “intuition” (or, at the very least, candidates thereof). In Section 7.4, I consider and reject the worry that my view trivially entails different meanings of “intuition” by examining the use of the word in Nagel (2012b). Then in Section 7.5 I defend the gap between my lexicographic account of Bengson and how Bengson himself theorizes about “intuition” by highlighting the gap between being competent at using language and being able to correctly theorize about language. Finally, in Section 7.6 I tie in this chapter’s discussion to themes present elsewhere in the dissertation.

7.2 Previous Metaphilosophical Analyses of “Intuition”

Starting before the rise of the contemporary intuition debate, van Inwagen (1997, 309) argues that “intuition” is nothing more than epistemic window-dressing. van Inwagen claims that “intuition” means nothing more than opinion, belief, or an inclination to believe.¹ In a brief passage, van Inwagen asserts that constructions like “intuition that P” are rhetorical slights-of-hand that lend claims more weight than “belief that P”.

¹ My focus here is primarily on “intuition” by people theorizing about “intuition”. While I have a few things to say about common English uses of “intuition” and run-of-the-mill philosophers’ use of “intuition”, my goal is to argue that “intuition” has technical polysemous meanings and that this may or may not be confined to philosophers who theorize about intuitions. This puts me slightly at odds with a few of the authors discussed in this section, especially van Inwagen and Williamson, so I flag where that difference matters.

Intuition-language is not about the unique phenomena of intuiting, van Inwagen claims, rather intuition-talk is a way to make our beliefs look like better support for claims than they are.

Building from this, Williamson (2007, 214-220) defends van Inwagen and a similar assertion by Lewis (1983, x). In contrast to van Inwagen, for whom discussion of intuition-talk is a passing point, Williamson devotes significant space trying to reverse-engineer what it is philosophers are talking about when they use “intuition”. Williamson is, in his own words looking at “clues to the role of the term “intuition” in contemporary analytic philosophy” (2007, 214). Among these clues are that the following contents are said to be the contents of intuitions:

- That the subject in a Gettier case lacks knowledge
- That there could be mountains
- That there are mountains (2007, 216, 219)

Williamson also cites an unnamed philosopher who said in the presence of Williamson that he had an intuition with the content *that they weigh more than three pounds* (2007, 214).²

Williamson runs through the different features of each of the mental states attached to these contents and compares them to rationalist accounts of intuitions, particularly the sui generis phenomenological account in Bealer (1998, 2002). Williamson finds that nothing substantial unites these mental states. *That there are mountains* is an inferential belief formed from empirical data, as is *I weigh more than three pounds*. Similarly *That Smith lacks knowledge* lacks any marked phenomenology,³ and, as Williamson argues elsewhere,

² These four data points are clearly meant to be metalinguistic data, but not all the data Williamson appeals to is clearly metalinguistic. It is somewhat difficult to tell where to demarcate reviews and criticism of the rationalist intuition literature from discussions of how “intuition” is used. Taking a broadly metalinguistic reading of Williamson’s discussion we can also add the linguistic data point that *intuitions are talked about as things that can be false* (2007, 216).

³ Or so Williamson testifies. Someone should look into whether the argument over whether intuitions have a special phenomenology is the byproduct of neurodiversity among metaphilosophers.

it is the product of fairly sophisticated modal inferences. So, Williamson concludes, there is little uniting these cases except that they are inclinations to believe, offering validity to van Inwagen’s observation.

Notice, however, that Williamson’s argument only works as stated if Williamson makes what I called in Chapter 5 the Univocal Assumption of Language. That is, Williamson assumes “intuition” only has one meaning. He takes the (by no means exhaustive) linguistic data he examines as evidence for a conclusion about the use of “intuition” simpliciter. Williamson assumes all philosophers mean the same thing (or intend to mean the same thing) by “intuition” in all contexts. By repeatedly criticizing Bealer’s account of intuitions for not capturing all uses of “intuition”, Williamson does not consider, for example, whether Bealer (1998) is using “intuition” – whether consciously or unconsciously – with a more restricted extension in mind than typically used by philosophers.⁴

If we assume for a moment that “intuition” is polysemous, then it is not surprising that Williamson came to a deflationary or thin reading of “intuition” analyzing the usage the way he did. Williamson moves from one observed usage of “intuition” to another, pointing out the ways in which the usages are collectively in tension with a thick reading of intuition-talk. Giving polysemous senses of words typically share some central theme, imagery, or function, if we run the same sort of analysis on uncontroversially polysemous words, we will often come up with a similarly thin reading.

To see this, consider polysemous senses of “bank” related to money-lending institutions.

One polysemous sense refers to the building:

(54) The coffee shop owners renovated an abandoned bank for their new location.⁵

⁴ Both Cappelen (2012, 57) and Devitt (2015, 681) grant that Bealer and BonJour (1998) are using “intuition” in a restricted sense.

⁵ The OED3 does not identify this as a unique sense of “bank”. Instead, the OED suggests – bizarrely in my eyes – this use is either obsolete or a shortened form of “bank building”. If you disagree that this is a unique sense of “bank”, feel free to ignore this example.

One sense refers to the corporate structure:

(55) I was nervous about switching to a bank without any physical branches, but it has worked out pretty well.

One sense refers to the branch of the corporate entity:

(56) Oh, you are with Royal Bank of Scotland? Where is their nearest bank?⁶

One sense refers to stores of something:

(57) At one point I was living off food from food banks and selling my plasma at a blood bank.

Finally, an American colloquial sense refers to large amounts of money:

(58) In her new job she’s making serious bank.

Imagine examining these different uses of “bank” while trying to determine what singular thing people mean by “bank”. With the Univocal Assumption, any answer will necessarily be very thin in order to accommodate all examples, especially, as Williamson does, if all uses are assumed to be literal and by language users competent in using the word. There is certainly shared themes in the uses of “bank”, but the uses do not share anything more robust than *related to storage or safekeeping*.⁷ Similarly, if there are multiple polysemous uses of “intuition”, trying to find a common meaning among them will likely lead to something as thin as *inclination to believe*.

In fact, when put through the sort of methodology Williamson uses, *inclination to believe* is a plausible reading of other epistemic terms that have multiple related uses.⁸

Consider the following uses of “perceived”:

⁶ You may want to combine this with the first sense. See previous footnote.

⁷ Even this definition over-generates, as we cannot call a safe in our bedroom or a squirrel’s acorn cache a bank.

⁸ See (Bengson, 2014, 560-562) for a related point in criticism of Cappelen (2012).

(59) I *perceived* the faint smell of burning rubber.

(60) I was trying to figure out why she was acting the way she does. After watching how she moved around the room, I *perceived* that she was agitated about her boyfriend.

(61) I *perceived* the force of his argument.

Only (59) is straightforwardly sensory. (60) is an inference based off of sensory data. (61) is remarking on a sort of internal moment of insight that is not obviously sensory, although undoubtedly many philosophers discussing experiences like that described in (61) will draw comparisons to sense data. Since all three uses of “perceive” discuss the acquisition of defeasible grounds for belief, someone looking for the shared meaning for all three may conclude that “A perceives X” means *A is inclined to believe X*, especially since even the straight-forwardly perceptual (59) does, at the end of the day, involve the speaker being inclined to believe that there is burning rubber. For this reason, any analysis of “intuition” – or indeed any word or phrase – needs to leave open the possibility that there is more than one meaning floating around.

7.2.1 Cappelen

Given the mistakes of Williamson’s analysis, we should not put much weight on it. Luckily he is not the only person to have examined philosophers’ use of “intuition”. Cappelen (2012) devotes almost half his book to a more sophisticated version of Williamson’s deflationary argument from philosophical usage of “intuition”, spending multiple chapters to examining lay usage of “intuition”, usage of “intuition” among philosophers, and more specialized usage among philosophers theorizing about intuitions. Here, Cappelen adds to Williamson’s approach by allowing for the possibility that “intuition” has different meanings across different subcommunities of English speakers. Cappelen considers two different ways the meaning may differ between subcommunities: between philosophers and

non-philosophers or between different groups of philosophers, such as between Kantians and Rawlsians. Cappelen thereby avoids the sort of worries I raise against Williamson. While the conclusion Cappelen draws about “intuition” is ultimately a bit muddled, he does nonetheless argue either that “intuition” is meaningless or that it means something like *an unreflective or pre-theoretic judgement* – even among metaphilosophers.

Cappelen’s arguments for the deflationary picture depend on the incorrect view that if specialized meanings of a word exist, use will differ along clear lines between subcommunities.⁹ While Cappelen considers the possibility that “intuition” has a special Philosophy-ese meaning, Cappelen makes a lot of hay out of philosophers’ use of “intuition” in the same way as non-philosophers (see Cappelen, 2012, 61-71). Passages from the likes of heavyweights such as Burge, Sider, Jaegwon Kim, and Williamson are shown to contain uses of intuition language that is either purportedly redundant or is used in the standard English meaning of, roughly, *arrived at easily and without deliberation*.

Credit goes to Cappelen here for doing lexicographical analysis of philosophical texts, but given the discussion in the previous chapters, finding non-specialized uses of “intuition” by philosophers is entirely unsurprising. Philosophers writing and speaking in English are after all writing and speaking in English. They are therefore going to use words or phrases as they would in standard English. If “intuition” has technical polysemous senses, then we should only expect some philosophers’ use of “intuition” to conform to the jargon meaning for the same reason that I only sometimes use “mouth” to refer to a feature of a river. Polysemy makes words very flexible, allowing me to mean many different things with a single morphological or phonological string. Sometimes one meaning of that string is suitable for my purposes, and sometimes a different meaning of that string is suitable for my purposes. In the case of polysemy jargon, the key difference is that I am privy of one sense of the word that most people outside of my subcommunity are not.

⁹ I get the sense Cappelen wants to make the stronger claim that certain facts about linguistic subcommunities ground different meanings of a word, but he does not say that outright.

Therefore, lexicographic evidence that philosophers use “intuition” in a standard English way, at least some of the time, is not evidence that “intuition” fails to be a theoretical term, at least some of the time.

Cappelen’s main argument against the existence of a specialized meaning of “intuition” comes from the level of disagreement among metaphilosophers working on intuitions. Specifically, Cappelen contrasts features of “intuition” with successful theoretical terms. He says that any theoretical term T is defective and does not successfully refer to anything of note if it has the following characteristics that “intuition” does (Cappelen, 2012, 51):

- T has no agreed upon definition among practitioners of the discipline.
- There is no agreement among participants in the discipline about what cases constitute core paradigms of the extension of T.
- There is no agreed upon theoretical role that T (or T’s extension) plays in the discipline.¹⁰
- There is considerable disagreement and dispute within the discipline about T itself.

Each of these observations plausibly describes intuition-talk, especially when we view the field of philosophy in the aggregate. However, as mentioned above, we should not view “intuition” as if it has a singular meaning, and Cappelen is open to the possibility that there are multiple meanings of “intuition”. If different meanings exist, Cappelen claims, they exist among subcommunities of speakers (2012, 57-60). Therefore, to tighten up Cappelen’s claims in light of my objections to Williamson’s argument about looking for singular meanings for intuition-talk among too wide of a linguistic net, we can adjust the above list to be about distinct philosophical subcommunities. Call this list **THE ADJUSTED CRITERIA**:

¹⁰ This observation is in tension with the no-theory theory of intuitions used in this dissertation, as this is one of the places Deutsch’s and Cappelen’s view differ significantly.

- “Intuition” has no agreed upon definition among members of philosophical subcommunities.
- There is no agreement among participants in the discipline (or subcommunities thereof) about what cases constitute core paradigms of the extension of “intuitions”.
- Among subcommunities of philosophers, there is no agreed upon theoretical role that “intuition” plays in the discipline.
- Among subcommunities of philosophers, there is no agreed upon extension of intuitions.
- There is considerable disagreement and dispute within subcommunities about intuitions themselves.

The Adjusted Criteria – which is just the original set of observations but about subcommunities – is not universally true about philosophical subcommunities. Kant scholarship is a particularly salient counterexample to the Adjusted Criteria, where use, extension, and understanding of the function of intuitions and “intuition” are tied to Kant’s texts.¹¹ Other plausible subcommunities that fail to meet the adjusted diagnostic criteria are Rawlsians, philosophers of cognitive science, and early analytics.¹² All of these subfields have specialized uses of “intuition” that seem to be dominant in the disciplines, at least as dominant as one would expect for a word with multiple polysemous senses. Nevertheless, I will set aside these counterexamples since Cappelen and myself are both interested in the way “intuition” has been used by the fields studied by post-DePaul and Ramsey (1998) metaphilosophy, which has primarily focused on the methodology of late 20th century analytic epistemology, with passing interests in late 20th century philosophy of language, philosophy of mind, metaphysics, and ethics.

¹¹ See Matthey (n.d.). Thanks to Stefano Lo Re and Janis Schaab for confirmation of this.

¹² Thanks to Daniel Healey for first bringing this to my attention as well as Saranga Sudarshan for help with Rawlsians.

The Adjusted Diagnostic Criteria brings Cappelen in line with one the main takeaways of last chapter:

FIRST LESSON OF JARGON: Jargon can belong to specific groups whose boundaries do not cleanly line up with disciplinary or sub-disciplinary boundaries.

The relevant barriers of polysemy jargon are not necessarily between philosopher and non-philosopher, but between Aristotelians and non-Aristotelians or between people who have learned the basic notions of formal semantics and those who have not. However, for reasons just discussed, if “intuition” is polysemous, we are not going to see unified usage of the word even within subfields. “Intuition”, besides the subcommunity-specific use (if it has them) is going to be used by members in that subcommunity in other ways, whether one of the standard English polysemous senses or the polysemy jargon senses of another subcommunity they belong to.

This does not in itself make disagreement among metaphilosophers less damning for “intuition”, since if “intuition” is polysemy jargon, members of the subcommunity should generally agree about the definition of the jargon sense of the word. While quantum physicists presumably agree with each other about “string”, metaphilosophers do not agree with each other about “intuition”. Quite the opposite is true, as illustrated by the following incomplete list of accounts of intuitions:

- Outputs of mental modules (Nagel, 2012b)
- Seemings (Bealer, 1998)
- Expressions of psychological concepts (Kornblith, 2015)
- Pre-theoretic judgements (Cappelen, 2012)
- Inclinations to believe (Williamson, 2007)

- Consciously inaccessible mental states Nado (2017)
- Intuitive presentations (Chudnoff, 2013b)
- Metacognitive feelings of ease (Egler, 2020)
- Beliefs (van Inwagen, 1997)
- Expressions of conceptual competence (Ludwig, 2007)
- Spontaneous judgements formed in reaction to cases (Weinberg et al., 2001)
- Judgements not inferred from beliefs (Kornblith, 1998)

My view is that some of these authors are putting forward candidate polysemy jargon meanings of “intuition”, and some may have succeeded in introducing (or at least propagating) polysemy jargon. Because they are introducing unique meanings of the word, specialized use will not be uniform across metaphilosophers, even after we remove non-specialised use of “intuition”. Instead, use will be uniform within and between certain papers, as metaphilosophers adopt and use one specific use for a paper’s purposes. Before I defend this, we need to introduce one last author, Devitt (2015), who defends a Belief Polysemy account of the intuition debate.

7.2.2 Devitt

Devitt (2015) is largely interested in defending the claim that intuitions play a key role in philosophy against Cappelen’s arguments that “intuition”, if not meaningless, means nothing more than *snap judgement*. In so doing, Devitt makes a few comments about the metasemantics of “intuition” relevant to this chapter’s discussion of metaphilosophers’ use of “intuition”. In the spirit of Kripke (1981) and Putnam (1975), Devitt lays out a Belief Polysemy account of the disagreement over intuitions by metaphilosophers. This is summed up in his rhetorical question: “What have these differences among theorists of

intuitions got to do with what they, let alone anyone else, mean by ‘intuition’?” (Devitt, 2015, 678). Disagreement in sciences over the extension or correct theory of trees, Devitt’s thought runs, is not a disagreement over the meaning of “trees”. Such disagreement is over the nature or essence of trees.¹³ Similarly, Devitt argues, disagreement revolving around the word “intuition” is just belief-level disagreement over the essence of intuitions.

Devitt does grant that some metaphilosophers use “intuition” as polysemy jargon – specifically Bealer (1998) and Pust (2000) who use “intuition” in a restricted sense to refer to necessary truths – but Devitt goes on to say

There is evidence that some metaphilosophers use ‘intuition’ with a special sense. But I doubt that there is evidence that most do and, in the absence of evidence, we should take the default position and assume that they don’t.
(2015, 681-682)

Since his attention is elsewhere, Devitt does not argue for why we should take this as the default position, but I take Devitt to be picking up on the default view semantic externalists (and philosophers more generally) have about debates in philosophy.¹⁴ Insofar as there’s a dispute centered around some contentious term “P”, philosophers assume disagreement boils down to philosophers’ different beliefs about P.

The last two chapters were written with an eye towards providing reasons why Belief Polysemy understandings of philosophical disputes like Devitt’s should *not* be the default assumption. I am not the first person to argue in an externalist framework that sometimes debates of the form “what is x” are really disputes over the actual or preferred meaning of

¹³ For reasons outlined in this dissertation, I do not think this is always true. As an example, at the time of writing, Wikipedia’s article on dinosaurs includes the line “The smallest known dinosaur is the bee hummingbird” (“Dinosaur”, 2020). Biologists now include birds as members of the dinosaur clade, so include birds among the referents of “dinosaur”. I take this to plausibly represent a new meaning of “dinosaur” that exists alongside the older one that refers to the things that died off 65 million years ago. See discussion of species in LaPorte (2009) for background on the topic. On the very slim chance anyone born after 2000 is reading this and was never taught dinosaurs all went extinct, I apologize for the nonsensical example.

¹⁴ See, however Chalmers (2011); Plunkett (2015) for notable rejections of this from internalist-leaning philosophers.

“x” (Sterken, 2020), although my framework is unique in that it does not set the different meanings in opposition to each other. This is because, just as in everyday language, meanings can co-exist with each other happily.

In Chapter 5, I discussed the ways in which language is generally polysemous and why those polysemous senses correspond to unique meanings of a word. Chapter 6 combined this discussion with specialized language, where I argued that in many cases technical notions or concepts can hide behind unassuming and common words. Specialists need to refer to distinctions or categories non-specialist language does not, so often common words are co-opted by specialists to refer to those distinctions. When we look at the way this happened with terms like “string”, “berry”, or “valid”, there is nothing mysterious here. The word is ambiguous, and many members of the general public do not know about the extent of their ambiguity because the ambiguity of these words were increased by specialists for the purpose of communicating amongst themselves.

If “intuition” is used in a variety of ways among metaphilosophers, then given discussion in the previous few chapters, we should suspect that “intuition” has branched meanings. We cannot determine this by looking at the debate from afar. As discussion of Williamson, Devitt, and Cappelen illustrate, incompatible uses of the same word can have multiple different semantic interpretations. Non-uniform use of a word in a corpus could have many causes, including:

- Total linguistic chaos where no idiolect agrees with any other idiolect on correct usage or meaning
- A single meaning with different metalinguistic beliefs
- A single meaning with different object-level beliefs
- Multiple meanings of a word spread universally among the population
- Multiple meanings of a word, some of which are limited to the subcommunity under

investigation

- Multiple meanings of a word, one of which is limited to a subset of the subcommunity under investigation
- Multiple meanings of a word, with multiple subcommunities having their own specialized meaning

Instead of listing off different accounts of “intuition” and using the fact *that there is widespread disagreement surrounding “intuition”* as the guiding evidence in our discussion of disagreement, we need to instead look at how different authors use “intuition” to find clues about what they mean. We have to do the hard work of lexicography.

7.3 “Intuition” as Polysemy Jargon

Two strategies have been employed in the previous attempts to analyze “intuition” in the texts described so far in this chapter. Although I only briefly discussed it, Cappelen does undertake lexicographic analysis, but his readings – at least to my eyes – require more context to be convincing (2012, 61-77). Moreover in these texts, Cappelen only finds either redundant and excisable or standard English uses of “intuition”. Absence of evidence is not evidence of absence, however, and this analysis only canvasses a handful of uses of intuition-talk in metaphysics and philosophy of language from the last 50 years. If there is any place where we would expect to find signs of polysemy jargon, especially if it is not obvious to philosophers, is in the most specialized texts on the topic. In this case, that is the intuition debate.

The far more common strategy, the one that does look at the intuition debate, is to view disagreement among philosophers from the abstract without careful examination of actual specific written examples. This lack of attention to actual texts is striking because Devitt’s Belief Polysemy view in particular entails certain predictions

about how metaphilosophers talk.¹⁵ If Meaning Polysemy is true for the intuition-debate, then we would expect this to be reflected in metaphilosophers’ discussion of intuitions. Metaphilosophers should either explicitly state their intent to capture the true nature of intuitions or at least remain quiet on the point, taking it as given that it would be understood that this is what they are doing when investigating intuitions. When looking at metaphilosophers’ discussion of intuitions, this is not always what we see.

7.3.1 Bengson

To start, consider the following passage from Bengson (2015), which is his response to the objection that his quasi-perceptual account of intuitions does not line up with either philosophers’ or everyday English’s use of “intuition”. Bengson writes:

[In] the present context, such sociological and linguistic concerns are largely beside the point. From our current perspective (i.e. the perspective of philosophical theorizing), what ultimately matters is not so much the contingent social role or ordinary use of a particular English term, but, first, whether there is a mental state of the sort described above and, second, whether such a mental state might serve as a legitimate epistemic source. We first located our target, not through sociological speculation or reflection on particular ordinary language locutions, but with examples. We have seen that the core quasi-perceptualist thesis accurately characterizes this target. Hence, in my view, the suggestion that a mental state that satisfies this characterization only imperfectly deserves the label “intuition” is merely terminological. Whether we choose to call it “intuition” or something else, such as “intellection” “insight”, or “quasi-perception”, it seems that once we have accurately

¹⁵ Cappelen (2012) makes other predictions about “intuition” beyond this chapter’s scope. See Andow (2015a, 2015b) for a quantitative corpus analysis that tests some of them.

characterized our target the philosophically interesting question is what work such a state might do (2015, 733-734).

Here, Bengson is clearly marking his use of “intuition” as polysemy jargon, albeit more than halfway through his 50-page paper. Importantly, Bengson clearly outlines his reasoning for introducing a polysemous sense of “intuition”. Bengson is interested in a certain category of mental states that he believes play an important theoretical role in the epistemology of philosophy – quasi-perceptual mental states – and there is no adequate label in English for this category. But, as he spends the whole paper arguing, the theoretical entities he is interested in, what he labels “intuitions”, go a long way in his mind to explain the phenomenology, epistemology, and psychology of considering thought experiments in philosophy and math. Other labels could have done the job, but he chose to use “intuition”.

It may seem strange that if Bengson is attempting to introduce a technical polysemous sense of “intuition” that he would make this intention clear earlier in the paper. However, Bengson does explain his specialized use earlier in the paper, albeit implicitly. Following the same pattern as many related works that ties intuitions to a phenomenal state (see, for example, Bealer (1998); Chudnoff (2013a); Koksvik (2017)), effort is spent early in the paper drawing the reader’s attention to a certain type of intellectual experience. Bengson in particular, devotes an entire section and five thought experiments to pin down what mental states intuitions are (2015, 709-715), distinguishing his intended target from other mental states, including other mental states that are sometimes called intuitions by philosophers. This is where Bengson educates readers on his specialized meaning of “intuition” demonstratively, although without stating his intention explicitly.

Whether or not Bengson has met the necessary and sufficient conditions for introducing polysemy jargon (I am purposefully remaining neutral on what these are), he is offering up a potential specialized meaning of “intuition” that goes above and beyond

normal or even philosophers’ meaning of “intuition”. He has a theoretical need to express something, and he does not have the expressive power from standard English or existing philosophical technical language. In order to give himself the expressive power (and undoubtedly to embed his discussion in the existing intuition debate), he borrows “intuition” and attempts to give it a new, specialized meaning.

7.3.2 Weinberg, Nichols, and Stich

Bengson is the clearest example in the literature of someone offering up a polysemous sense of “intuition”, but other less-clear examples exist as well. Consider the following passage in Weinberg et al. (2001) (hereafter WNS), one of the first and most influential works of experimental philosophy:

As we use the notion, an epistemic intuition is simply a spontaneous judgment about the epistemic properties of some specific case – a judgment for which the person making the judgment may be able to offer no plausible justification (2001, 19)

WNS are marking their use of “intuition” as a restricted sense of the sort of standard English meanings of “intuition” discussed above, specifically limiting them to spontaneous judgements *about a case* and *in reaction to a case*. This excludes spontaneous judgements about principles (such as that the axiom of choice holds) or spontaneous judgements about mathematical truths, when these are not elicited by thought experiments – putting WNS’s use at odds with Bengson’s, as well as other metaphilosophers with views similar to Bengson (Bealer, 1998; Chudnoff, 2017; Egler, 2020).

The illocutionary force of WNS’s passage is less clear-cut than Bengson’s, since WNS could either be following Bengson in stipulating a special meaning of “intuition” or disambiguating between existing senses of “intuition”. Due to restrictions of time and space,

historical lexicography in this chapter is kept to a minimum (see Hintikka, 1999), and I am not in a position to comment on whether they are disambiguating between uses of “intuition” that were in the literature prior to 2001. That said, a case can probably be made that they are following the technical polysemous sense of “intuition” that often appears in the heuristics and biases literature as well as the related dual-processing literature (see Gendler, 2010). This disambiguation reading already assumes “intuition” has technical polysemous senses, even if philosophers are sharing jargon with social psychologists and cognitive scientists (see again the First Lesson of Jargon from last chapter).

Focusing on the possibility that WNS are stipulating new polysemy jargon, thinking about what WNS are trying to accomplish in Weinberg et al. (2001), the need for stipulating new jargon is clear. WNS were at the time using a novel experimental method to question the evidential value of using thought experiments in rationalist analytic epistemology. Arguing from survey results about thought experiments, the paper claims that philosophical reasoning from thought experiments may be the product of acculturation. Given the central methodological and argumentative importance of survey answers, which recorded (or at least attempted to record) participants’ spontaneous reactions, WNS needed some way of talking about their data points. Since they were talking about a subset of intuitions (understood in the standard English spontaneous-judgement sense) *and* intuition-language is prominent in philosophical discussion of thought experiments, a restricted or metonymic use of “intuition” was a natural name for the target notion. WNS had a theoretical need to talk about an category of mental states that did not have a name, and so they co-opted “intuition”. This is well within their right as specialists.

It is worth pointing out as well that, whether they stipulated it or borrowed an existing meaning, WNS’s use of “intuition” stuck among experimental philosophers. Experimental philosophers until recently loved using “intuitions” to talk about their data points. Speaking anecdotally, recent criticism of intuition-language from people like Deutsch and

Cappelen and responses to such criticisms that drop intuition-language – such as Colaço and Machery (2017) – have led to experimental philosophers dropping intuition-language in favor of language about judgments.

7.4 Objection 1: Return of the Semantic Anarchy Worry

In relationship to sense-talk in the previous chapter, I raised and dismissed the worry that my view leads to semantic anarchy where words have arbitrarily high numbers of meanings (Section 6.3.4). Now that I have introduced two purported technical polysemous meanings of “intuition”, I can discuss the specific worry that my methodology entails all metaphilosophers writing on intuitions have their own (attempted) semantic meaning of “intuition”. By granting philosophers like WNS and Bengson power to create their own meaning of “intuition”, the objection runs, my view is committed to one or both of the following TRIVIAL POLYSEMY SCHEMA for any philosopher P:

- If P defends or assumes that intuitions are essentially Q, then “intuition” has Q as a polysemous meaning.
- If P defends a unique account of intuitions, then they have introduced a new polysemous meaning of “intuition”.

Devitt accuses Cappelen of a similar mistake. Devitt argues Cappelen was misled by semantic ascent from intuition to “intuition” into thinking “intuition” is meaningless (2015, 679). One might worry I am falling into a similar trap. I do not think Bengson and WNS exhaust the ways in which philosophers have given “intuition” new technical polysemous senses – I take there to also be a separate etiological meaning of “intuition” in Sosa (2007) and Ludwig (2007) – but there is a limit to how many senses of “intuition” are floating around metaphilosophy. As a matter of fact, I think experimental

philosophers have, in general, stuck to either the everyday spontaneous-sense of “intuition” or WNS’s restricted spontaneous-reaction-to-case-sense of “intuition”. Similarly, I take the phenomenologists I cited in relationship to Bengson (Bealer (1998); Chudnoff (2013a); Koksvik (2017)) to be using “intuition” in the same way and offering different *accounts* of the same category of mental state. They are giving competing accounts of the phenomenological state that gives us the feeling of truthiness in relationship to some proposition when we are thinking about math and philosophy (if such a state exists).

To defuse the Semantic Anarchy Worry, it is important to pay attention to my methodology. Borrowing from lexicography, I am examining specific uses of “intuition” in early 21st century metaphilosophy. Back in Section 6.3.5, I explained how lexicographers have evidence from context, use, and their own linguistic judgements at their disposal, and in this case, I combined context clues from Bengson (2015) and Weinberg et al. (2001) such as the author’s stated intentions and theoretical background, usage of “intuition” by that author, and my own judgements about usage and context as an English speaker. Separating out senses lexicographically is more of an art than a science, and leave it to the critical reader to offer their own lexicographic analyses of these authors to counter the claims here. Nevertheless, the inferences here are more sophisticated than plugging works of metaphilosophy into either of the Trivial Polysemy Schema.

7.4.1 Response: Nagel’s Use of “intuition”

To drive home the point that this method does not commit me to a different meaning of “intuition” for every paper or author writing on intuitions, I will examine Nagel (2012b). Despite appearances to the contrary, Nagel is using “intuition” to mean the same thing as WNS, although she has a specific theory of the etiology of such states.

Nagel places her discussion of intuitions within the context of the dual-systems tradi-

tion in cognitive science.¹⁶ Very roughly speaking, some of our cognition is done by slow, effortful, and conscious cognitive architecture, and some of our thinking is done by fast, easy, and unconscious cognitive architecture (J. S. B. T. Evans & Frankish, 2009). The quick system is then sometimes argued to be made up of modules, which are discrete, domain-specific systems that are informationally insulated from the rest of the brain. Nagel defends the claim that epistemic intuitions are outputs of a mindreading module – a module responsible for interpreting the mental states of other people. She then draws conclusions on the reliability of philosophical intuitions based on empirical data on the epistemic features of other modular inferences.

Put at this level of generality, it seems then I am committed to something like the following:

Nagel’s polysemy-jargon use of “intuition” means *judgements that arise out of cognitive architecture that is fast, quick, and easy (i.e., mental modules)*, or Nagel’s use of “intuition” is an attempt to establish the polysemous meaning thereof.

This is not the case. When we dig deep into the text of Nagel (2012b) for clues about “intuition”, we find Nagel is using “intuition” in line with WNS above.

One set of clues comes in the way she quotes and follows experimental philosophers’ use of “intuition” without comment. For example, she quotes a use of “intuition” by a paper co-authored by Weinberg (of WNS fame) and spends pages discussing the findings of WNS as being about intuitions without indicating she takes herself to mean something else by the word. Consider the following quote from when Nagel is motivating the importance of thinking about the epistemology of the mental states studied by experimental philosophers:

¹⁶ This is a slight oversimplification. Nagel’s main source is Mercier and Sperber (2009), who defend a massively modular view of the mind. I conflate massive modularity and dual systems theory here for ease of presentation.

Intuitions may exhibit worrisome instability either within an individual (Swain et al., 2008, 335), or between groups of individuals, perhaps along such epistemically scary fault lines as ethnicity (Weinberg et al., 2001) or gender (Stich & Buckwalter, 2011). (Nagel, 2012b, 495)

Nagel is making no effort to distance herself from the use of “intuition” in those experimental studies.

Contrast that passage by Nagel with a nearly identical passage where Ludwig discusses the experimental findings of Machery et al. (2004):

The results indicate a significant difference between the Western participants and the East Asian participants in responses to the Gödel cases, with Western participants significantly more often giving (B) answers. The differences in *responses* to the Jonah cases, however, were not statistically significant. [emphasis added] (Ludwig, 2007, 136)

Instead of following the standard intuition-talk, Ludwig replaces “intuition” with “response”. This is not the only way Ludwig distances himself from experimental philosophers’ intuition-talk. In another paper, he writes:

Clearly [my] use of “intuition” to mean intellectual intuition is distinct from the stipulative use of “intuition” specified by WNS. (2010, 437)

Ludwig, while talking about WNS, takes WNS’s use of “intuition” to be stipulated and thus different than how he uses “intuition”. Ludwig takes his own use of “intuition” to be the proper, pre-existing meaning of the word (more on this below in Section 7.6.1). The above two quotes thus reveals Ludwig’s metalinguistic beliefs that WNS’s use of “intuition” does not track his own. Such passages are absent in Nagel’s work, indicating she is using the word in the same way as WNS.

A second set of clues indicating Nagel intends for “intuition” to be in line with WNS, and not cognitive states with a certain etiology, is that when Nagel is interested in talking about states with a certain cognitive etiology, she uses a word other than “intuition”. Throughout the paper, Nagel draws a careful distinction between intuitions and intuitive judgements. Nagel explicitly borrows the terminology of “intuitive judgement” from Mercier and Sperber (2009) to contrast with “reflective judgements” (2012b, 498) as different sorts of judgments coming from different types of cognition. The use of these two different noun-phrases and the clunkiness of “intuitive judgement” compared to “intuition” indicates she sees a distinction between the two.

Building on this distinction, a third set of clues comes from the argumentative structure of Nagel’s paper. “Intuition” is not *defined* in her paper in terms of intuitive judgements, rather Nagel *defends* the claim that the reactions we have about cases are a subset of intuitive judgements. To see this, compare two sentences that appear almost side-by-side. The first sentence is:

There are no obvious barriers to seeing epistemic case intuitions as the products of our ordinary mindreading capacities. (2012b, 521)

At this point in the paper, Nagel is concluding her argument that the reactions people have to philosophical thought experiments have the same psychological and epistemic features as ordinary mindreading capacities. The claim is that judgments I have about knowledge reading Gettier cases goes through the same cognitive processes as my everyday judgments about, say, whether my partner knows I spilled wine on the living room carpet.

Pair that quote with Nagel’s claim one paragraph later, keeping in mind Nagel is borrowing “intuitive” from Mercier and Sperber (2009):

Mindreading, like perception and other intuitive capacities, is susceptible to certain natural illusions [...] (2012b, 521)

This quote indicates that Nagel takes mindreading judgments to have a specific cognitive origin, namely some sort of quick-and-dirty system dedicated to forming epistemic judgments (Mercier and Sperber (2009) for their part hold a massive modularity view of the mind). She is defending the substantial thesis that intuitions are intuitive judgments, by which she is following WNS in using “intuition” to mean reactions to thought experiments and Mercier and Sperber that “intuitive” judgements are inferences with certain cognitive properties.

This somewhat belabored discussion of Nagel is to show that my account of polysemy jargon in philosophy does not overgenerate and is not committed to anything like the Trivial Polysemy Schema. Instead, the claim here is that a handful of theorists have found it useful to co-opt “intuition” to mean something slightly different than its everyday English meaning, and that it is well within those theorists’ power and right to do so. WNS wanted a term to talk about the mental states captured by their experiments and Bengson wanted a term to talk about mental states with a certain phenomenology. Insofar as their usage has gained currency among (meta)philosophers, then we should think they succeeded in creating new meanings of “intuition”, albeit ones with very limited circulation. We should therefore not think, as Devitt does, that “intuition” is unified in meaning and referent or that, as Williamson and Cappelen do, that a lack of a clear unified pattern of the use of “intuition” is evidence that “intuition” univocally means something like *a belief* or *a pre-theoretic judgment*. Like natural language in general, the language of philosophy is messy and full of words with multiple related meanings, and we need to do lexicographical analysis of philosophers’ use of language to discover what these are.

7.5 Objection 2: Metalinguistic Claims in Bengson (2014)

In a different paper to the one discussed above, Bengson (2014) defends an account of the intuition debate that gives reasons for thinking Bengson is not intending to stipulate a new meaning of “intuition.” Bengson (2014) responds to Cappelen’s arguments from intuition-language (which I covered parts of in Section 7.2.1) by arguing that metaphilosophers use “intuition” in neither the everyday way or a technical way. Instead, when metaphilosophers use “intuition”, Bengson argues, they use a *discriminatory* sense of “intuition” to pick out a theoretically relevant subset of intuitions (2014, 559-565). Metaphilosophers, Bengson’s argument runs, are not stipulating a new use of “intuition”. Instead metaphilosophers are using “intuition” to denote a proper subset of what “intuition” normally denotes, and this is not the same as a technical meaning of the word.

If Bengson’s views did not change between his 2014 response to Cappelen and his 2015 discussion of intuitions, then it seems I am misinterpreting Bengson’s use of “intuition” in the passage quoted above. Indeed, just before the above quoted passage, Bengson calls normal intuition-talk “far less discriminate” than his own (2015, 733). So perhaps Bengson is not stipulating a new term at all, but rather using “intuition” in a normal way, but more discriminatorily.

However, Bengson’s view about “intuition” being a normal but discriminatory use in the mouths of metaphilosophers is betrayed by the grammar of his own intuition-talk. In English, we *can* use words in a more discriminatory sense without new senses of a word forming, but it leaves certain syntactic markers. To see this let’s consider two nouns, “apple” and “couch”, which like “intuition” have subsets and can be used either as count nouns about specific tokens (“the couch in my living room”, “some apples fell off the tree”, “an intuition I had about the case”) or in the plural form about the corresponding kind

(“intuitions are difficult to theorize about”, “apples are generally sweet”, “The Romans invented couches”).

To give Bengson as much benefit of the doubt as possible, imagine we were in a slightly different context, where I was writing on the history of Gala apples. I could write:

(62) The apples were first bred in New Zealand.

Here I am using “apples” in this discriminatory way Bengson is supposing. I want to talk about a relevant subset of apples, Gala apples, and I did so with “apples”. It is important to note that the full noun phrase is “the apples”, though, and dropping “the” ends up changing the meaning to something false (apples were first domesticated in ancient Central Asia):

(63) Apples were first bred in New Zealand.

So while “the apples” referentially picks out Gala apples through context ((62)), simply “apples” ((63)) does not.

Similarly, imagine I was instead writing a dissertation on sectional couches:

(64) The couches’ modular design fits any space.

(65) Those couches’ modular design fits any space.

(66) Couches’ modular design fits any space.

(67) A couch’s modular design fits any space.

Both (64) and (65), which include “the” and “those” both allow for discriminate use of “couch” to refer to sectional couches. However, unmodified, “couches” does not, nor does “a couch” in (66) and (67). So, we can refer to a subset of some kind in the discriminatory way Bengson supposes, but not when the word is unmodified and lacks “the”, “those” or some other grammatical unit.

This is not how we see Bengson or other metaphilosophers talk about intuitions. Bengson does not write “those intuitions x” or “the intuitions x”. Instead Bengson writes:

Rather, the point is that there are varieties of non-voluntariness, and presentational states such as experiences and intuitions are non-voluntary in a way that imaginings, guesses, hypotheses, beliefs, or judgements are not. (2015, 722)

and:

It is argued, first, that *intuitions* and perceptual experiences are at a certain level of abstraction the same type of mental state, presentations, which are distinct from beliefs, hunches, inclinations, attractions, and seemings. [emphasis added] (2015, 709).

The syntactic structure of Bengson’s use of “intuition” does not support Bengson’s discriminatory metasemantic theory. Adding to my point, Bengson also uses “intuition” as a singular collective noun:

Presentationism, if true, provides an explanation of the justificatory status of *intuition*.” [emphasis added] (2015, 744)

Lexicographers tend to identify changes in syntax like this – from count noun to singular collective noun – as evidence of different polysemous senses (Atkins & Rundell, 2008).

It appears Bengson, in writing about metaphilosophers’ use of “intuition”, misunderstands how metonymic shifts in language work. Metonymic shifts are very common mechanisms behind polysemy, where a word gains a second meaning that is closely related in some way to the original meaning, and pretty much every case of polysemy discussed in the dissertation was the product of some sort of metonymic branching or another (see Cruse, 2000, 109-120). Bengson is trying to square the way metonymy occurs in technical

settings with the Univocal Assumption. This leads him to argue that “intuition” maintains one meaning, albeit one that can be modulated based on the needs of the author. As I have argued, we have strong reasons to think the Univocal Assumption is false and that the relationship between word and meaning is plastic and typically a 1-to-many relationship instead of typically a 1-to-1 relationship. So, while Bengson theorizes that his work is a restriction of “intuition”, it is in fact the use or formation of a technical polysemous sense.

7.5.1 Theorizing About Language Vs Using Language

Notice I am committed to two claims that might appear to be in tension. First, I observed that Bengson theorizes that his use of “intuition” is the same as normal English. Bengson (2014) argues many metaphilosophers’ intuition-talk is just a discriminatory use that denotes a proper subset of the normal use of “intuition”. Second, I argued that Bengson (2015) is stipulating a new meaning of “intuition”, although I suggested in discussion of the semantic anarchy worry (Section 7.4) that he may in fact be using a pre-existing polysemy jargon sense used by phenomenologists. A similar tension also appears above in my response to the semantic anarchy worry in Section 7.4, where, after arguing at the start of this chapter and in Chapter 5 that philosophers falsely hold the Univocal Assumption, I nonetheless argue that Nagel and Ludwig competently navigate polysemy around “intuition”. These claims, despite the apparent tension, are compatible because philosophers can employ certain language and be wrong when they theorize about their language. Philosophers can be better at using language than theorizing about it.

Even setting aside arguments from semantic externalism allowing for false beliefs about a word’s meaning, this disconnection between explicit and implicit understanding of language should be familiar to anyone reading. It is a strange feature of human language processing, but a feature nonetheless, that we can be competent users of language without

being good at explaining our linguistic choices. Anyone has experienced this phenomena if they have tried to explain to a non-native speaker one of the trickier features of English grammar – such the differences between count and non-count nouns or between “which” and “that” – or if they have used a word in a sentence but struggled seconds later to define the word. Therefore, there is no problem in saying that a competent user of English such as Bengson theorizes in one paper that he and his colleagues are using a discriminatory use of “intuition”, while actually using the word in another way.

Theorizing about language is hard and being good at speaking a language does not mean you can automatically give a correct theory of what you are doing. This is why I can argue that Bengson (2014) is not a definitive guide to the metalinguistics of Bengson (2015) or that philosophers can be wrong about the polysemy of “intuition” but still swap between polysemous senses of it.

7.6 Moving the Intuition Debate Forward

To finish this chapter and the dissertation, it is time to look at the consequences of this discussion of the polysemy jargon of “intuition”. The above discussion has demonstrated that we should not assume that there is one answer to the issue of what intuitions are. Many different theorists mean different things by “intuition” because they are interested in slightly different things involving the case method. Among the others discussed, WNS and Nagel are interested in the spontaneous reactions people have about cases while Bengson is interested in mental states that carry a certain presentation of truth. It is unfortunate they (and others) co-opted “intuition” to talk about these, but the choice is understandable. When philosophers unreflectively talk about intuitions involving the case method, they are generally talking about reactive propositional attitudes with significant doxastic force or epistemic value. The spirit of this is retained to some degree in all the polysemy jargon discussed in this chapter.

From this, we have two important upshots I look at in turn. First, accounts of intuitions are not criticisable on the grounds they involve some non-standard meaning of “intuition.” In fact, metaphilosophers should feel free to play around with what they mean by “intuition” as long as they are clear with what they are doing. Second, and returning to Chapter 1, the no-theory theory of intuition should be used with ample caution. Because there are different meanings of “intuition” that differ between competing theorists, the no-theory theory is ripe for misunderstanding.

7.6.1 Non-Standard Usage of Terms is Fine

The language in this chapter and the last has been purposefully ecumenical. While I think metaphilosophers could be more explicit about their semantic intentions – if they are even aware of them – I do not think there is any harm in adjusting language to serve one’s needs. To explore this ecumenical view, I will push back against Ludwig’s (2010) criticism of WNS, which rests partly on the claim that WNS misuse “intuition”.

Ludwig is admirably clear about his metalinguistic intentions.¹⁷ He explicitly intends for his use of “intuition” to have the same meaning it has when philosophers discuss the case method (2010, 437). Ludwig argues that a priori philosophical inquiry is ultimately after conceptual truths, and we employ thought experiments to discover these conceptual truths. Since philosophers have this goal and method in mind when discussing intuitions about thought experiments (or so Ludwig claims) the correct understanding and meaning of intuitions are judgments that are the products of our competence deploying (Fregean) concepts in thought. In other words, Ludwig argues my mental state with the content *that Smith lacks knowledge* is only an intuition if I had the mental state because I formed it on the basis of correctly deploying the concept KNOWLEDGE in thought. It is not an intuition on Ludwig’s understanding if I form the mental state on the basis of a memory,

¹⁷ This chapter (indeed the entire second half of the dissertation) was largely inspired by Ludwig (2007, 2010).

hunch, derive it from a belief, or use some other heuristic that does not draw competently upon the concept KNOWLEDGE.

As quoted above, Ludwig is willing to grant that WNS at least tried to stipulate a new meaning of “intuition” (2010, 437), and Ludwig similarly considers other “candidates for intuitions” and “classifications” of intuitions (2010, 435). He rejects these, however, on the grounds that these are not the sorts of things philosophers have been interested in when using intuitions and discussing intuitions.¹⁸ This is a variation on Devitt’s point above, where Ludwig is arguing that there is a single correct meaning of “intuition” and any variation from this is a mistake. In contrast to Devitt who is mostly interested in metasemantics, Ludwig is primarily interested in conceptual issues and thus willing to call out other uses of “intuition” as conceptually foolish. E.g.,

No competent philosopher, in instructing someone in what the point of a philosophical thought experiment is, would say to her that it was to get her to make a spontaneous judgment about it. That would be absurd.

Ludwig takes the concept INTUITION to be inextricably tied in with the nature of philosophy and the use of thought experiments, so pulling those apart by stipulating a new meaning of “intuition” involves deep conceptual mistakes about the nature of INTUITION. Ludwig’s argument ultimately rests on the teleological point that we use thought experiments to inform our theories about conceptual truths, and, more importantly for present purposes, the conceptual point about INTUITION that intuitions are the mental states that put us in touch with those conceptual truths. Any metalinguistic move that draws away from this is a mistake.

Hopefully my view here is obvious. Learning conceptual truths via thought experiments is not the only reason people use intuition-talk. WNS were not interested in conceptual truths, rather they were interested in the ways reactions to cases differed along

¹⁸ Note Ludwig’s claim is not: what philosophers have been interested in when using “intuition”.

cultural and socioeconomic lines. Insofar as WNS were interested in conceptual truths, they were interested in the *sociological* phenomena of using reflective equilibrium to try to get to them. Similarly, Nagel was interested in the reliability of the psychological processes that give rise to reactions to thought experiments. Bengson was in contrast interested in developing and explicating how certain things seeming true to us at a point in time leads to justification and knowledge. In this way, WNS and Bengson were engaging in what some would consider conceptual engineering (Burgess, Cappelen, & Plunkett, 2020), as they are building new concepts in order to better theorize about certain phenomena. Assuming then that Ludwig is right about the nature of the concept INTUITION, because of the metalinguistic moves of WNS, Bengson, and others, the concepts SPONTANEOUS-REACTION-TO-CASE and QUASI-PERCEPTION are now attached to the word “intuition” too.

Philosophers are not always going to be interested in the same set of questions, even when connected closely in subject matter. We should encourage them to experiment in theoretical and conceptual space, which may involve “misusing” intuition-talk. While philosophers should clarify metasemantic and conceptual issues when they do this, the practice should not be written off or denigrated.¹⁹

7.6.2 The No-Theory Theory is Risky

Returning to a topic discussed in Section 1.4.1, we now have a worry for no-theory theories of intuitions, such as those employed by Deutsch (2015) and Malmgren (2011). To reiterate, the no-theory theory of intuitions is the strategy of side-stepping contentious

¹⁹ A similar argument can be made against the genealogical arguments found in Hintikka (1999). Hintikka argues against the use of intuitions based on the claim that none of the methodological claims behind past uses of intuitions by theorists such as Chomsky, Plato, Aristotle, Descartes are or should be accepted by modern philosophers. While not the only line of attack against the use of intuitions in Hintikka (1999), hopefully my response to this sort of genealogical argument is obvious. Despite the joint use of “intuition” across these different philosophers (or translations thereof), we should not assume that they are all talking about the same things. For Hintikka’s argument to work, he therefore needs to argue that “intuition” means the same thing across all these authors.

claims about intuitions by not taking a stand on what intuitions are. By relying on a producer and consumer's theory-neutral understanding of intuitions, people employing the no-theory theory (including myself in the present work) aim to make claims about intuitions while avoiding tangential debates about the content, etiology, and phenomenology of intuitions.

Given discussion in this chapter, a flaw arises for the no-theory theory. Because there are multiple technical meanings or candidate technical meanings for "intuition" in metaphilosophical literature, any advantage the neutrality the no-theory theory affords is potentially offset by a high risk of misinterpretation. We cannot assume readers of a work employing the no-theory theory will all be on the same page with the author or each other when reading sentences involving intuition-talk. To see this, consider how Deutsch introduces the no-theory theory (note that, contrary to Nagel, Deutsch is using "intuitive judgement" and "intuition" interchangeably) (Deutsch, 2015, 24-25):

I have offered no theory of intuitions – none, at any rate, meant to reveal the nature or essence of intuitions.

In part, this is because I did not have to. In each of the cases I discussed, it was clear what the intuitive judgment relative to the case under discussion was. In the Gettier cases, it was, in the Weinberg et al. 2001 study, the judgment *that Bob really knows that Jill owns a Ford*, and, in the Starmans and Friedman 2012 study, it was the judgment *that Peter really knows that there is a watch on the table*. In the Machery et al. 2004 study on Gödel cases, it was the judgment *that John uses "Gödel" to talk about the man who stole the proof*. In the Knobe 2003 study, it was the judgment *that the chairman intentionally helped/harmed the environment*. And so on. And those are just a few of the intuitive judgments that xphiles have tested.

Deutsch then says of examples like this:

The term “these judgments” in the previous sentence has a reasonably clear reference. Call those judgments “intuitions.” “Are intuitions used as evidence?” or “What is the role of intuitions in philosophical argument?” now have reasonably clear senses and may be fruitfully discussed. No theory of intuition-hood is required. (Deutsch, 2015, 25)

On one reading, Deutsch is using “intuition” in some standard way, and all he is doing is not giving an account of the referent of the standard use of “intuition”. This is essentially Devitt’s view of the debate, where “intuition” means something and disagreement is merely over beliefs related to the meaning. As discussed in relationship to Ludwig, we might doubt whether there is one category of things here to talk about, or we may doubt that philosophers are interested in the referent of normal uses of “intuition”.

Another charitable reading is that Deutsch, like WNS and Bengson, is stipulating a bespoke meaning of “intuition” for his theoretical purposes. On this second reading, Deutsch is remaining as neutral as possible by providing an extensional meaning of “intuition”. In the passage quoted above and in one other passage (2015, 25), Deutsch points to seven mental states, and this may be the term’s stipulated extension or attempted stipulated extension. While perhaps initially less plausible than the first reading, Deutsch does not need to mean anything more for his exegesis-based intuition denial. Combining the claim that *those things* are not used as evidence for philosophical positions, and that *those things* are uncontroversially intuitions, he can inferentially reject intuitions’ evidential role in philosophy.²⁰

Using “intuition” in either of these two ways, while being neutral, increases the likelihood of misunderstanding. Deutsch (2015) was published after Weinberg et al. (2001) stipulated their meaning, after Bealer (1998) used “intuition” in a way restricted to certain types of seemings, and after Ludwig (2010) argued “intuition” could only mean

²⁰ Thanks to Derek Ball for this suggested reading of Deutsch.

mental states that are the products of conceptual competence. Each of these philosophers used “intuition” in a particular way because of their theoretical background and aims. This means that strings or words like “Are intuitions used as evidence?” are ambiguous between different meanings of “intuition”, including common meanings and technical meanings. By not taking a stance, Deutsch does not provide the interpretive clues to enable us as readers disambiguate the meaning of the sentence between existing senses. This risks bleeding into how readers interpret and understand the work. I, for one, suspect I default to a WNS-style spontaneous-reaction-to-case reading of “intuition”, and I suspect I have read and reread parts of Deutsch’s book and papers with that interpretation. Similarly, since we seem to interpret polysemous words in whatever way is the most common for us (Giora, 2003), other authors with different theoretical backgrounds used to using “intuition” in a certain way will plausibly default to interpreting sentences according to their own use.

Demonstratives will not fix the ambiguity. Pointing to the judgement *that Bob really knows that Jill owns a Ford* allows for the exactly the same sort of confusion. Is the mental state attached to that content a spontaneous judgment in reaction to a case? Is it a quasi-perceptual mental state? Is it the product of conceptual competence? Is it an inclination to believe? Each of these different views of intuition roughly agree on the content of the intuitions involving the states, although there may be some quibbling over whether the judgments have modal content. But despite roughly agreeing on the content, they disagree fundamentally about the properties of the mental state with the content. Because “intuition” is ambiguous, the no-theory theory will again exacerbate any misunderstanding as it leaves it open for each consumer to read the work as using their preferred meaning of “intuition”.

I am to some degree guilty as well since I use the no-theory theory throughout the first four chapters of the dissertation. This might have caused regrettable confusion, but

I aimed to make sure everything I said about intuitions came out true on all views of intuitions. When I argued in Chapters 3 and 4 that exegesis is an unmotivated method, intuition-language merely illustrated the standard intuition-friendly account of the case method against exegesis-based intuition denial.

Intuition-talk in Chapter 2 is a bit different since it assumes that intuitions (regardless of what the word means) have a certain function. My argument in Chapter 2 is conditional – if we use original texts to argue that thought experiments do not involve intuitions, then we are left with a skepticism about whether philosophers form beliefs based on good arguments. Here, the no-theory theory is advantageous. I am not interested in any account of intuitions. Rather I am interested in the functional role intuitions are thought to play as ways of justifying verdicts about thought experiments. To avoid skepticism, we need some good reason to take the thought experiment as having the verdict. So insofar as I discuss intuitions, I am only committed to them being reactive attitudes of some sort having the functional role that they carry some sort of epistemic weight related to thought experiment verdicts. Since the argument is conditional, I am not even committed to the existence of such mental states. Rather, my claim is that without anything playing that epistemic role (whatever that role is), the strategy of defusing experimental-driven skepticism by examining original texts will not work.

7.7 Conclusion: Beyond One Meaning

In this chapter, I argued that “intuition” has developed multiple meanings when used by metaphilosophers working on issues surrounding the case method. In particular, I rejected previous metalinguistic discussions of the intuition-debate by Williamson, Cappelen, Devitt, and Ludwig in favor of a more ecumenical understanding of how language in philosophy works. Bengson means “intuition” in a certain special way, and WNS and Nagel mean “intuition” in a different special way. This does not mean intuition-talk is

meaningless or that there is one privileged use of “intuition”. Rather, language is evolving. For the same reason we need to be careful about how we as listeners interpret or as speakers contextualize ambiguous sentences like “I see a star”, “He bought a mint”, or “these peppers are very hot”, we need to be careful about sentences about intuitions.

This chapter’s survey of the technical polysemous senses of “intuition” was not meant to be exhaustive. Rather, I discussed the authors I discussed in order to illustrate how the framework of polysemy jargon outlined in previous chapters applies to the intuition-debate and to offer a few lessons about how the intuition debate should move forward. Most metaphilosophers are not as clear about their metalinguistic intentions as Ludwig, WNS, and Bengson. In this way, Nagel’s tacit deference to WNS and experimental philosophers is more the norm than Bengson’s explicit writing-off of everyday meanings of “intuition”. More lexicography needs to be done if we want to understand the true scope of polysemy jargon.

This chapter is not meant to suggest “intuition” is the only term in philosophy that a lexicographic analysis would reveal has technical polysemous senses. I am confident, for example, analysis of “concept” would reveal much the same thing as found for intuition-talk. Like intuition-talk, concept-talk has a loose everyday use and significantly divergent uses in philosophical settings. But divergent uses between philosophers is not the only clue we have of polysemy jargon. When looking at a specific word, semantic and syntactic differences go hand-in-hand (Atkins & Rundell, 2008, 294-297). This means we should suspect polysemy jargon when philosophers use words with a different syntax than normal people.

To pick one set of examples, philosophers often use words as count-nouns that are not typically used as count-nouns in everyday discourse. Philosophers often use both “perception” and “belief” (or their conjugations) as count-nouns even though they are rarely used as count-nouns in everyday discussion. Whereas in everyday discussion I might

“believe in Santa Claus” and “perceive a faint odor”, philosophers say I “have a belief that Santa Claus exists” and “have a perception as of a faint smell”. In cases like these, even if philosophers speak in a unified way, we have reason to suspect they are using the word with a different meaning than non-philosophers.

With the framework developed in the last few chapters, philosophers will therefore be better equipped to spot non-standard meanings of a word and adjust their theorizing accordingly.

Chapter 8

Final Thoughts

Befitting a work on metaphilosophy, this dissertation defended two main theses – one object-level and one methodological. My object-level thesis is negative. The method of using original texts in order to defuse the experimental attack is flawed on multiple grounds. I argued in Chapter 2 that the method is self-defeating because it fails to take into account the historical and sociological fact that philosophers have changed their beliefs because of thought experiments. If we remove intuitions from works of thought experiments, as Deutsch and Cappelen argue we should, then we are left with bad arguments. Crucially, these are bad arguments philosophers have changed their minds because of. In Chapter 3 and Chapter 4, I argued that the method of looking at original texts in the first place is unmotivated. When we look at the epistemology of consuming philosophical texts, we can see that most knowledge caused by a work of philosophy is based on the philosophical judgments of the consumer, not producer, of the work. Therefore, there is little reason to examine original texts for actual methods, as opposed to checking consumers for what methods they used or thought the author used when reading the texts.

My methodological thesis is positive. While Deutsch and Cappelen back the wrong horse, they are right in spirit that epistemologists of philosophy need to explore new

methods. The armchair and experimental methods currently at metaphilosophers' disposal cannot answer all the questions metaphilosophers have asked. Chapters 2, 3, and 4 argued that metaphilosophers need to develop a rich research program that studies how philosophers interact with other philosophers. This will require drawing upon methods in anthropology, sociology, and developing novel methods in psychology. If we want to know about the epistemic states of people reading works of philosophy, we need to know the history of how those works reverberated around philosophy as well as the assumptions and thought processes of the philosophers reading those works.

Chapters 5, 6, and 7 drew upon the methods of applied linguistics, specifically lexicography. The lexicography of philosophy is not new, although past attempts to lexicographically analyze "intuition" have been too cursory or relied on bad metasemantic assumptions. Moreover, lexicography is often denigrated in philosophical circles as merely capable of discovering word use or word meaning. To push against this, I used lexicography of "intuition" to demonstrate that lexicography is a powerful and underutilized metaphilosophical tool. Philosophers often engage in sophisticated arguments that center around words that are also common in everyday discussion (e.g., good, right, know, reason, possible). Given the tendency of words to develop multiple related meanings (Chapter 5) including in technical settings (Chapter 6), it is an open question whether this has happened in philosophical debates as well (Chapter 7).

8.1 Future Directions

No completed PhD dissertation is a perfect dissertation, and this is a completed dissertation.¹ There are therefore many directions for future research.

One salient direction for future work is to actually engage in the anthropological,

¹ This syllogism is not sound at the time of writing. Hopefully it will not be sound in a different way at the time of reading.

sociological, and psychological projects I propose at the end of my discussion of how philosophers learn from each other. While I would love to engage in these in the future, I lack the relevant expertise in all three disciplines. I suspect starting the sociological and anthropological projects in particular will face considerable practical hurdles. Anthropology – and to a lesser extent sociology – draw upon works of philosophy in their theorizing, but the philosophy they draw upon is from the so-called continental tradition. This places a serious theoretical and linguistic barrier between analytic philosophy and the discipline of anthropology and sociology. Moreover, doing the sort of non-psychological project well would require getting an experienced anthropologist or sociologist on board. This requires convincing an anthropologist or sociologist that such a project would be worth their time. Given how inside baseball this dissertation’s suggested project is, this will be a tall order. It is nonetheless an order we should try to make.

Another particularly salient direction for future work is to complete the arc between the epistemology of philosophy that started the dissertation and the philosophy of language that ended it.² Discussion of “intuition” in the previous chapter has most of the pieces in place to argue that the intuition-debate is a merely verbal dispute, especially on less demanding accounts of verbal disputes (e.g., Inga, 2018). However, I am not convinced it is a merely verbal dispute. Having read influential works in the debate with an eye towards the metalinguistics of “intuition”, I cannot shake sense that participants in the debate understand that different camps use “intuition” with different (candidate) meanings. Therefore, the debate better resembles a metalinguistic negotiation between competing jargon senses of “intuition” (Plunkett, 2015), where different authors are tacitly putting forth accounts as the best (as opposed to actual) meaning of “intuition”.

Even if the intuition debate is not a verbal dispute or metalinguistic negotiation, the lessons that can be drawn from the jargon polysemy of “intuition” are complex and not fully fleshed out here. In one sense, it is a cautionary tale about polysemy jargon.

² This was planned but cut short by COVID-19.

Criticisms of intuition-language tracing back from van Inwagen through Williamson and Cappelen correctly observe that intuition-talk carries an extremely high level of rhetorical weight among philosophers. For that reason, “intuition” is a prime target for metaphilosophers to co-opt for theoretical purposes. Teasing this thought out, however, requires an in-depth exploration of the role of lexical effects in language design.

Any lessons drawn from the multiple meanings of “intuition” will be ecumenical. Different theorists have used the word in different ways because there are lots of different strands to explore involving the psychology and epistemology of thought experiments. Given they are all interested in slightly different issues, these projects all seem worthwhile. Phenomenologists like Bengson are interested in why and how certain internally identifiable mental states give us reasons to believe things. Empirically minded philosophers like WNS and Nagel are interested in what psychology and cognitive science can tell us about our epistemic standing. Ludwig is interested in connecting this all up with truth. To take part in each of these projects, we need to talk about different categories. While so far “intuition” has been used differently by all these projects, moving forward metaphilosophers should be clearer with themselves and their readers that intuitions are not always intuitions.

8.2 The End

Future directions aside, the takeaway of this work is hopefully clear. The subject matter of philosophy and the discipline of philosophy are necessarily intertwined. The practice of philosophy is done by humans interacting with other humans. The rise of experimental philosophy has offered a powerful tool for investigating how our contingent psychology affects our belief-formation, but this is only the beginning. Epistemologists of philosophy need to look past the individual philosopher and study the discipline of philosophy in all of its interpersonal and historically influenced messiness.

References

- Abbott, B. (1997). A Note on the Nature of "Water". *Mind*, 106(422), 311–319.
- Adeberg, T., Thompson, M., & Nahmias, E. (2015, July). Do men and women have different philosophical intuitions? Further data. *Philosophical Psychology*, 28(5), 615–641. doi: 10.1080/09515089.2013.878834
- Alexander, J., & Weinberg, J. M. (2007). Analytic epistemology and experimental philosophy. *Philosophy Compass*, 2(1), 56–80.
- Andow, J. (2015a, November). How Distinctive Is Philosophers' Intuition Talk? *Metaphilosophy*, 46(4-5), 515–538. doi: 10.1111/meta.12151
- Andow, J. (2015b, April). How "Intuition" Exploded. *Metaphilosophy*, 46(2), 189–212. doi: 10.1111/meta.12127
- Andow, J. (2016a). Abduction by Philosophers: Reorienting Philosophical Methodology. *Metaphilosophy*, 47(3), 353–370.
- Andow, J. (2016b, March). Thin, Fine and with Sensitivity: A Metamethodology of Intuitions. *Review of Philosophy and Psychology*, 7(1), 105–125. doi: 10.1007/s13164-015-0247-2
- Andow, J. (2017, January). Intuition-Talk: Virus or Virtue? *Philosophia*. doi: 10.1007/s11406-016-9796-6
- Andow, J. (2018, December). Why Don't Philosophers Do Their Intuition Practice? *Acta Analytica*. doi: 10.1007/s12136-018-0377-0
- Anscombe, E. (1979). What is it to Believe Someone? In C. F. Delaney (Ed.), *Rationality and Religious Belief*. University of Notre Dame Press.
- Ashton, Z., & Mizrahi, M. (2018). Intuition Talk is Not Methodologically Cheap: Empirically Testing the "Received Wisdom" About Armchair Philosophy. *Erkenntnis*, 83(3), 595–612. doi: 10.1007/s10670-017-9904-4
- Atkins, B. T. S., & Rundell, M. (2008). *The Oxford guide to practical lexicography*. Oxford ; New York: Oxford University Press.
- Audi, R. (1997). The Place of Testimony in the Fabric of Knowledge and Justification. *American Philosophical Quarterly*, 34(4), 405–422.
- Ball, D. (2020). Relativism, Metasemantics, and the Future. *Inquiry: An Interdisciplinary Journal of Philosophy*, 41.
- Barrett, H. C. (2020a, September). Deciding what to observe: Thoughts for a post-WEIRD generation. *Evolution and Human Behavior*, 41(5), 445–453. doi: 10.1016/j.evolhumbehav.2020.05.006
- Barrett, H. C. (2020b, August). Towards a Cognitive Science of the Human: Cross-

- Cultural Approaches and Their Urgency. *Trends in Cognitive Sciences*, 24(8), 620–638. doi: 10.1016/j.tics.2020.05.007
- Bartels, D. M., & Pizarro, D. A. (2011, October). The mismeasure of morals: Antisocial personality traits predict utilitarian responses to moral dilemmas. *Cognition*, 121(1), 154–161. doi: 10.1016/j.cognition.2011.05.010
- Baz, A. (2016, January). Recent Attempts to Defend the Philosophical Method of Cases and the Linguistic (Re)turn. *Philosophy and Phenomenological Research*, 92(1), 105–130. doi: 10.1111/phpr.12106
- Bealer, G. (1996). On the possibility of philosophical knowledge. *Philosophical Perspectives*, 10, 1–34.
- Bealer, G. (1998). Intuition and the Autonomy of Philosophy. In M. DePaul & W. Ramsey (Eds.), *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry* (pp. 201–240). Lanham, Md.: Rowman & Littlefield.
- Bealer, G. (2002). Modal Epistemology and the Rationalist Renaissance. In T. S. Gendler & J. Hawthorne (Eds.), *Conceivability and Possibility*. Oxford University Press.
- Beg the Question*. (2016). <https://www.merriam-webster.com/words-at-play/beg-the-question>.
- Belleri, D. (2018). Two Species of Merely Verbal Disputes. *Metaphilosophy*, 49(5), 691–710. doi: 10.1111/meta.12335
- Bengson, J. (2014). How philosophers use intuition and 'intuition'. *Philosophical Studies*, 171(3), 555–576.
- Bengson, J. (2015). The Intellectual Given. *Mind*, 124(495), 707–760.
- Bird, A. (2007, March). What Is Scientific Progress? *Nous*, 41(1), 64–89. doi: 10.1111/j.1468-0068.2007.00638.x
- Bloom, P. (2007). Water as an artifact kind. *Creations of the mind: Theories of artifacts and their representation*, 150–156.
- Boghossian, P. (1996). What the Sokal Hoax Ought to Teach Us. *Times Literary Supplement*.
- BonJour, L. (1998). *In Defense of Pure Reason: A Rationalist Account of A Priori Justification*. Cambridge: Cambridge University Press.
- Bourget, D., & Chalmers, D. J. (2014). What Do Philosophers Believe? *Philosophical Studies*, 170(3), 465–500.
- Boyd, K., & Nagel, J. (2014). The Reliability of Epistemic Intuitions. In E. Machery & O. Elizabeth (Eds.), *Current Controversies in Experimental Philosophy* (pp. 109–127). Routledge.
- Brown, J. (2017). The Gettier Case and Intuition. *Explaining knowledge: New essays on the gettier problem*, 191–211.
- Buckwalter, W., & Stich, S. (2010, September). *Gender and Philosophical Intuition* (SSRN Scholarly Paper No. ID 1683066). Rochester, NY: Social Science Research Network. doi: 10.2139/ssrn.1683066
- Bullock, O. M., Colón Amill, D., Shulman, H. C., & Dixon, G. N. (2019, July). Jargon as a barrier to effective science communication: Evidence from metacognition. *Public Understanding of Science*, 0963662519865687. doi: 10.1177/0963662519865687

- Burge, T. (1998). Computer proof, apriori knowledge, and other minds. *Noûs*, 32(S12), 1–37.
- Burgess, A., Cappelen, H., & Plunkett, D. (Eds.). (2020). *Conceptual engineering and conceptual ethics* (First edition ed.). Oxford: Oxford University Press.
- Cappelen, H. (2012). *Philosophy without intuitions*. Oxford: Oxford University Press.
- Cappelen, H. (2014a). Replies to Weatherson, Chalmers, Weinberg, and Bengson. *Philosophical Studies*, 171(3), 577–600.
- Cappelen, H. (2014b). X-Phi without Intuitions? In A. R. Booth & D. P. Rowbottom (Eds.), *Intuitions*. Oxford University Press.
- Cappelen, H. (2018). *Fixing Language*. Oxford University Press. doi: 10.1093/oso/9780198814719.001.0001
- Carey, S. (2011). *The origin of concepts* (1. iss. Oxford Univ. paperback ed.). Oxford: Oxford Univ. Press.
- Castro, C. M., Wilson, C., Wang, F., & Schillinger, D. (2007, September). Babel babble: Physicians’ use of unclarified medical jargon with patients. *AMERICAN JOURNAL OF HEALTH BEHAVIOR*, 31, S85-S95.
- Casullo, A. (2003). *Two Conceptions of A Priori Justification*. Oxford University Press.
- Chalmers, D. J. (1996). *The conscious mind: In search of a fundamental theory*. Oxford university press.
- Chalmers, D. J. (2011). Verbal disputes. *Philosophical Review*, 120(4), 515–566.
- Chalmers, D. J. (2014). Intuitions in philosophy: A minimal defense. *Philosophical Studies*, 171(3), 535–544.
- Chalmers, D. J. (2015). Why Isn’t There More Progress in Philosophy? *Philosophy*, 90(1), 3–31.
- Chudnoff, E. (2011, March). The nature of intuitive justification. *Philosophical Studies*, 153(2), 313–333. doi: 10.1007/s11098-010-9495-2
- Chudnoff, E. (2013a). *Intuition* (First edition ed.). Oxford, United Kingdom ; New York, NY: Oxford University Press.
- Chudnoff, E. (2013b). Intuitive knowledge. *Philosophical Studies*, 162(2), 359–378.
- Chudnoff, E. (2017). The reality of the intuitive. *Inquiry*, 60(4), 371–385. doi: 10.1080/0020174X.2016.1220640
- Chudnoff, E. (in press). Mature Intuition. *Synthese*.
Cleve Crudginton (No. 34). (2005, November). BBC Two.
- Climenhaga, N. (2017). Intuitions are Used as Evidence in Philosophy. *Mind*, fzw032.
- Coady, C. A. (1992). *Testimony: A philosophical study*.
- Cokely, E. T., & Feltz, A. (2009, February). Individual differences, judgment biases, and theory-of-mind: Deconstructing the intentional action side effect asymmetry. *Journal of Research in Personality*, 43(1), 18–24. doi: 10.1016/j.jrp.2008.10.007
- Colaço, D., & Machery, E. (2017, May). The intuitive is a red herring. *Inquiry*, 60(4), 403–419. doi: 10.1080/0020174X.2016.1220638
- Constantin, J., & Grundmann, T. (2018, September). Epistemic authority: Preemption through source sensitive defeat. *Synthese*. doi: 10.1007/s11229-018-01923-x
- Cook, M. (2018, October). *Hoax offers hope that post-modernist madness is*

- fading*. <https://www.mercatornet.com/features/view/hoax-offers-hope-that-post-modernist-madness-is-fading/21811>.
- Cooper, R. (2007). *Psychiatry and philosophy of science*. Stockfield: Acumen.
- Cova, F., Strickland, B., Abatista, A., Allard, A., Andow, J., Attie, M., . . . Zhou, X. (2018). Estimating the Reproducibility of Experimental Philosophy. *Review of Philosophy and Psychology*.
- Cruse, D. A. (2000). *Meaning in language: An introduction to semantics and pragmatics*. Oxford: Oxford Univ. Press.
- Cullen, S. (2010). Survey-Driven Romanticism. *Review of Philosophy and Psychology*, 1(2), 275–296.
- Cummins, R. C. (1998). Reflection on Reflective Equilibrium. In M. DePaul & W. Ramsey (Eds.), *Rethinking Intuition* (pp. 113–128). Lanham, Md.: Rowman & Littlefield.
- Darwin, C., & Bynum, W. F. (2011). *On the origin of species by means of natural selection: Or the preservation of favoured races in the struggle for life*. Cambridge; London: ProQuest LLC ; Penguin.
- Davidson, D. (1987). Knowing one’s own mind. *Proceedings and Addresses of the American Philosophical Association*, 60(3), 441–458.
- DePaul, M. R., & Ramsey, W. M. (Eds.). (1998). *Rethinking intuition: The psychology of intuition and its role in philosophical inquiry*. Lanham, Md: Rowman & Littlefield.
- Deutsch, M. (2009). Experimental Philosophy and the Theory of Reference. *Mind & Language*, 24(4), 445–466. doi: 10.1111/j.1468-0017.2009.01370.x
- Deutsch, M. (2010). Intuitions, counter-examples, and experimental philosophy. *Review of Philosophy and Psychology*, 1(3), 447–460.
- Deutsch, M. (2015). *The Myth of the Intuitive: Experimental Philosophy and Philosophical Method*. Cambridge, Massachusetts: The MIT Press, a Bradford Book.
- Deutsch, M. (2016). Gettier’s Method. In J. Nado & J. R. Beebe (Eds.), *Advances in Experimental Philosophy and Philosophical Methodology* (pp. 69–98). London ; New York: Bloomsbury Academic.
- Devitt, M. (2011). Experimental Semantics. *Philosophy and Phenomenological Research*, 82(2), 418–435.
- Devitt, M. (2013). Is There a Place for Truth-Conditional Pragmatics? *Teorema*, 32(2).
- Devitt, M. (2015, November). Relying on Intuitions: Where Cappelen and Deutsch Go Wrong. *Inquiry*, 58(7-8), 669–699. doi: 10.1080/0020174X.2015.1084824
- Dinosaur. (2020, September). *Wikipedia*.
- Dorr, C., & Hawthorne, J. (2014). Semantic plasticity and speech reports. *Philosophical Review*, 123(3), 281–338.
- Dutant, J. (2015). The legend of the justified true belief analysis. *Philosophical Perspectives*, 29(1), 95–145.
- Ebert, P. A., Smith, M., & Durbach, I. (2018, January). Lottery judgments: A philosophical and experimental study. *Philosophical Psychology*, 31(1), 110–138. doi: 10.1080/09515089.2017.1367767
- Egler, M. (2019, November). No hope for the Irrelevance Claim. *Philosophical Studies*. doi: 10.1007/s11098-019-01373-2

- Egler, M. (2020, February). Testing for the phenomenal: Intuition, metacognition, and philosophical methodology. *Mind & Language*, 35(1), 48–66. doi: 10.1111/mila.12229
- Egler, M., & Ross, L. D. (2020). Philosophical expertise under the microscope. *Synthese*, 197(3), 1077–1098.
- Evans, G. (1973). The Causal Theory of Names. *Aristotelian Society Supplementary Volume*, 47(1), 187–208.
- Evans, J. S. B. T., & Frankish, K. (Eds.). (2009). *In two minds: Dual processes and beyond*. Oxford ; New York: Oxford University Press.
- Everett, T. J. (2014). Peer Disagreement and Two Principles of Rational Belief. *Australasian Journal of Philosophy*, 93(2), 273–286.
- Fischer, E., & Engelhardt, P. E. (2017a). Diagnostic Experimental Philosophy. *Teorema: International Journal of Philosophy*, 36(3), 117–137.
- Fischer, E., & Engelhardt, P. E. (2017b). Stereotypical Inferences: Philosophical Relevance and Psycholinguistic Toolkit. *Ratio*, 30(4), 411–442.
- Fischer, E., Engelhardt, P. E., & Herbelot, A. (2015). Intuitions and illusions: From explanation and experiment to assessment. In E. Fischer & J. Collins (Eds.), *Experimental Philosophy, Rationalism, and Naturalism. Rethinking Philosophical Method* (pp. 259–292). Routledge.
- Fischer, E., Engelhardt, P. E., Horvath, J., & Ohtani, H. (in press). Experimental Ordinary Language Philosophy: A Cross-Linguistic Study of Defeasible Default Inferences. *Synthese*, 1–42.
- Foot, P. (1967). The problem of abortion and the doctrine of double effect. *Oxford Review*, 5, 5–15.
- Frege, G. (1948). Sense and reference. *The philosophical review*, 57(3), 209–230.
- Frege, G. (1997). Logic (P. Long & R. White, Trans.). In M. Beaney (Ed.), *The Frege Reader* (pp. 227–250). Blackwell Oxford.
- Fricker, E. (2014). Others—An Argument from Analogy? *Religious faith and intellectual virtue*, 174.
- Fuller, S. (2002, June). Prolegomena to a Sociology of Philosophy in the Twentieth-Century English-Speaking World. *Philosophy of the Social Sciences*, 32(2), 151–177. doi: 10.1177/004931032002002
- Geeraerts, D. (1993, January). Vagueness’s puzzles, polysemy’s vagaries. *Cognitive Linguistics*, 4(3), 223–272. doi: 10.1515/cogl.1993.4.3.223
- Gendler, T. S. (2000). *Thought experiment: On the Powers and Limits of Imaginary Cases*. New York: Routledge.
- Gendler, T. S. (2010). *Intuition, imagination, and philosophical methodology*. Oxford: Oxford University Press.
- Gert, B., & Gert, J. (2017). The Definition of Morality. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2017 ed.). Metaphysics Research Lab, Stanford University.
- Gettier, E. (1963). Is Justified True Belief Knowledge? *Analysis*, 23(6), 121–123.
- Giora, R. (2003). *On our mind: Salience, context, and figurative language*. New York

- Oxford: Oxford University Press.
- Goldberg, S. (2013, May). Epistemic Dependence in Testimonial Belief, in the Classroom and Beyond. *Journal of Philosophy of Education*, 47(2), 168–186. doi: 10.1111/1467-9752.12019
- Goldman, A. I. (1999). *Speech Regulation and the Marketplace of Ideas*. Oxford University Press.
- Goldman, A. I. (2010). Philosophical naturalism and intuitional methodology. *Proceedings and Addresses of the American Philosophical Association*.
- Goodman, N. (1983). *Fact, Fiction, and Forecast* (4th ed ed.). Cambridge, Mass: Harvard University Press.
- Google Books Ngram Viewer of "rhapsody". (2020). <https://books.google.com/ngrams/graph?content=rhapsody>.
- Greco, J. (2016, December). What is Transmission*? *Episteme*, 13(4), 481–498. doi: 10.1017/epi.2016.25
- Grey, W. (2000). Metaphor and meaning. *Minerva: An Internet Journal of Philosophy*, 4.
- Grice, H. P. (1957, July). Meaning. *The Philosophical Review*, 66(3), 377. doi: 10.2307/2182440
- Grice, H. P. (1993). *Studies in the way of words* (3. print ed.). Cambridge, Mass.: Harvard Univ. Press.
- Grundmann, T. (2013). Doubts About Philosophy? The Alleged Challenge From Disagreement. In T. Henning & D. Schweikard (Eds.), *Knowledge, Virtue, and Action. Essays on Putting Epistemic Virtues to Work* (pp. 72–98). Routledge.
- Hanlon, A. (2019, January). *How blind reverence for science obscures real problems*. <https://www.washingtonpost.com/outlook/2019/01/28/how-blind-reverence-science-obscures-real-problems/>.
- Hannikainen, I. R., Machery, E., & Cushman, F. A. (2018, January). Is utilitarian sacrifice becoming more morally permissible? *Cognition*, 170, 95–101. doi: 10.1016/j.cognition.2017.09.013
- Hansen, N., Porter, J. D., & Francis, K. (2019). A Corpus Study of "Know": On the Verification of Philosophers' Frequency Claims about Language. *Episteme*, 1–27.
- Hawley, K. (2010, December). Testimony and knowing how. *Studies in History and Philosophy of Science Part A*, 41(4), 397–404. doi: 10.1016/j.shpsa.2010.10.005
- Hazlett, A. (2007). Grice's Razor. *Metaphilosophy*, 38(5), 669–690. doi: 10.1111/j.1467-9973.2007.00512.x
- Helzer, E. G., & Pizarro, D. A. (2011, April). Dirty Liberals!: Reminders of Physical Cleanliness Influence Moral and Political Attitudes. *Psychological Science*, 22(4), 517–522. doi: 10.1177/0956797611402514
- Henderson, D. K., & Horgan, T. (2013). *The Epistemological Spectrum: At the interface of cognitive science and conceptual analysis*. Oxford: Oxford Univ. Press.
- Hills, A. (2009, October). Moral Testimony and Moral Epistemology. *Ethics*, 120(1), 94–127. doi: 10.1086/648610
- Hills, A. (2020). Moral Testimony: Transmission Versus Propagation. *Philosophy and*

- Phenomenological Research*. doi: 10.1111/phpr.12595
- Hintikka, J. (1999, March). The Emperor's New Intuitions. *The Journal of Philosophy*, 96(3), 127–147. doi: 10.5840/jphil199996331
- Horvath, J. (2010, August). How (not) to react to experimental philosophy. *Philosophical Psychology*, 23(4), 447–480. doi: 10.1080/09515089.2010.505878
- Hu, M., & Nation, P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a foreign language*, 13(1), 403–30.
- Huemer, M. (2011, August). Epistemological Egoism and Agent-Centered Norms. In T. Dougherty (Ed.), *Evidentialism and its Discontents* (pp. 17–33). Oxford University Press. doi: 10.1093/acprof:oso/9780199563500.003.0002
- Hume, D. (1748). *Of Miracles*.
- Ichikawa, J. J. (2016). Intuitive Evidence and Experimental Philosophy. In J. Nado & J. R. Beebe (Eds.), *Advances in Experimental Philosophy and Philosophical Methodology* (pp. 155–174). London ; New York: Bloomsbury Academic.
- Ichikawa, J. J., & Jarvis, B. W. (2013). *The rules of thought* (1st ed.). Oxford: Oxford University Press.
- Indo-European Lexicon: PIE Etyma and IE Reflexes*. (n.d.). <https://lrc.la.utexas.edu/lex>.
- Inga, V. (2018). Verbal Disputes and the Varieties of Verbalness. *Erkenntnis*, 83(2), 331–348. doi: 10.1007/s10670-017-9892-4
- Jenkins, C. S. I. (2014). Merely Verbal Disputes. *Erkenntnis*, 79(S1), 11–30.
- Kahneman, D. (2013). *Thinking, fast and slow* (1st pbk. ed ed.). New York: Farrar, Straus and Giroux.
- Kaplan, D. (1989). Demonstratives: An Essay on the Semantics, Logic, Metaphysics and Epistemology of Demonstratives and other Indexicals. In J. Almog, J. Perry, & H. Wettstein (Eds.), *Themes From Kaplan* (pp. 481–563). Oxford University Press.
- Kaplan, M. (1985). It's not what you know that counts. *The Journal of Philosophy*, 82(7), 350–363.
- Kauppinen, A. (2007, June). The Rise and Fall of Experimental Philosophy. *Philosophical Explorations*, 10(2), 95–118. doi: 10.1080/13869790701305871
- Kelly, T. (2005). The epistemic significance of disagreement. *Oxford studies in epistemology*, 1(167-196).
- Klein, D. E., & Murphy, G. L. (2001, August). The Representation of Polysemous Words. *Journal of Memory and Language*, 45(2), 259–282. doi: 10.1006/jmla.2001.2779
- Knowles, E. (2010). *How to read a word*. Oxford ; New York: Oxford University Press.
- Koksvik, O. (2017, January). The phenomenology of intuition. *Philosophy Compass*, 12(1), e12387. doi: 10.1111/phc3.12387
- Kornblith, H. (1998). The role of intuition in philosophical inquiry: An account with no unnatural ingredients.
- Kornblith, H. (2010, August). Belief in the Face of Controversy. In R. Feldman & T. A. Warfield (Eds.), *Disagreement* (pp. 29–52). Oxford University Press. doi: 10.1093/acprof:oso/9780199226078.003.0003

- Kornblith, H. (2012). *On reflection* (1st ed.). Oxford: Oxford University Press.
- Kornblith, H. (2013). Is Philosophical Knowledge Possible? In D. E. Machuca (Ed.), *Disagreement and Skepticism* (p. 260). Routledge.
- Kornblith, H. (2015). Naturalistic Defenses of Intuition. In E. Fischer (Ed.), *Experimental philosophy, rationalism, and naturalism: Rethinking philosophical method* (First ed.). New York: Routledge.
- Kosem, I. (2016, March). Interrogating a Corpus. In P. Durkin (Ed.), *The Oxford Handbook of Lexicography*.
- Kripke, S. (1981). *Naming and Necessity*. Oxford, UK ; Cambridge, USA: Blackwell Publishers.
- Kuhn, T. (1977). A function for thought experiments.
- Kuhn, T. (1996). *The structure of scientific revolutions* (3rd ed ed.). Chicago, IL: University of Chicago Press.
- Kuroki, G. (2000, September). *After The Sokal Affair and Impostures Intellectuelles*. <http://hps.elte.hu/~gk/Sokal/Sokal/Linkek/Kuroki-Sokal.html>.
- Kusch, M. (2002). *Knowledge by agreement: The programme of communitarian epistemology*. Oxford : New York: Clarendon Press ; Oxford University Press.
- Lackey, J. (2008). *Learning from Words*. Oxford University Press. doi: 10.1093/acprof:oso/9780199219162.001.0001
- Ladyman, J., Ross, D., Spurrett, D., & Collier, J. G. (2007). *Everything must go: Metaphysics naturalized*. Oxford ; New York: Oxford University Press.
- Lakoff, G. (1987, June). The Death of Dead Metaphor. *Metaphor and Symbolic Activity*, 2(2), 143–147.
- Lakoff, G., & Johnson, M. (1980). *Metaphors We Live By*. University of Chicago Press.
- Landes, E. (2020, March). The threat of the intuition-shaped hole. *Inquiry*, 1–26. doi: 10.1080/0020174X.2020.1742784
- Landes, E. (ms). Designing the Language of Pandemic: COVID-19, "Coronavirus", and "Social Distancing".
- LaPorte, J. (2009). *Natural kinds and conceptual change*. Cambridge University Press.
- Lehrer, K. (1990). *Theory of Knowledge*. London: Routledge.
- Lewis, D. K. (1983). *Philosophical papers*. New York: Oxford University Press.
- Liao, S. M., Wiegmann, A., Alexander, J., & Vong, G. (2012, October). Putting the trolley in order: Experimental philosophy and the loop case. *Philosophical Psychology*, 25(5), 661–671. doi: 10.1080/09515089.2011.627536
- Libben, G., Gibson, M., Yoon, Y. B., & Sandra, D. (2003, January). Compound fracture: The role of semantic transparency and morphological headedness. *Brain and Language*, 84(1), 50–64. doi: 10.1016/S0093-934X(02)00520-5
- Links, A. R., Callon, W., Wasserman, C., Walsh, J., Beach, M. C., & Boss, E. F. (2019, June). Surgeon use of medical jargon with parents in the outpatient setting. *Patient Education and Counseling*, 102(6), 1111–1118. doi: 10.1016/j.pec.2019.02.002
- Ludlow, P. (2014). *Living words: Meaning underdetermination and the dynamic lexicon*. OUP Oxford.
- Ludwig, K. (2007). The epistemology of thought experiments: First person versus third

- person approaches. *Midwest Studies in Philosophy*, 31(1), 128–159.
- Ludwig, K. (2010, August). Intuitions and relativity. *Philosophical Psychology*, 23(4), 427–445. doi: 10.1080/09515089.2010.505877
- Ludwig, K. (2013). Methods in Analytic Epistemology. In M. C. Haug (Ed.), *Philosophical methodology: The armchair or the laboratory?* (First ed.). New York: Routledge.
- Lycan, W. (2019). Representational Theories of Consciousness. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2019 ed.). Metaphysics Research Lab, Stanford University.
- Machery, E. (2017). *Philosophy within its proper bounds* (First ed.). Oxford, United Kingdom: Oxford University Press.
- Machery, E., Mallon, R., Nichols, S., & Stich, S. P. (2004). Semantics, cross-cultural style. *Cognition*, 92(3), 1–12.
- Machery, E., & Seppälä, S. (2011). Against hybrid theories of concepts. *Anthropology and Philosophy*, 10, 99–126.
- Machery, E., Stich, S. P., Rose, D., Chatterjee, A., Karasawa, K., Struchiner, N., ... Hashimoto, T. (2018). Gettier was framed. *Epistemology for the Rest of the World*, 123–148.
- Maitra, I. (2018, September). New Words for Old Wrongs. *Episteme*, 15(3), 345–362. doi: 10.1017/epi.2018.27
- Malmgren, A.-S. (2011, January). Rationalism and the Content of Intuitive Judgements. *Mind*, 120(478), 263–327. doi: 10.1093/mind/fzr039
- Mattey, G. J. (n.d.). *G. J. Mattey's Kant Lexicon: Intuition*. hume.ucdavis.edu/mattey/phi175/intuition.html.
- Mercier, H., & Sperber, D. (2009). Intuitive and Reflective Inferences. In *In Two Minds: Dual Processes and Beyond* (pp. 149–168). New York: Oxford University Press.
- Meskin, A., & Liao, S.-y. (2018). Experimental Philosophical Aesthetics as Public Philosophy. In S. Réhault & F. Cova (Eds.), *Advances in Experimental Philosophy of Aesthetics* (pp. 309–326). Bloomsbury.
- Millar, R. M., & Trask, R. L. (Eds.). (2015). *Trask's historical linguistics* (3rd Edition ed.). New York: Routledge.
- Misak, C., & Talisse, R. B. (2019, November). *Pragmatism Endures*. <https://aeon.co/essays/pragmatism-is-one-of-the-most-successful-idioms-in-philosophy>.
- Mizrahi, M. (in press). Conceptions of Scientific Progress in Scientific Practice: An Empirical Study. *Synthese*.
- Moran, R. (2006, June). Getting Told and Being Believed. In J. Lackey & E. Sosa (Eds.), *The Epistemology of Testimony*. Oxford University Press. doi: 10.1093/acprof:oso/9780199276011.001.0001
- Mortensen, K., & Nagel, J. (2016). Armchair-friendly experimental philosophy. *A companion to experimental philosophy*, 53–70.
- Nadelhoffer, T., & Feltz, A. (2008, July). The Actor–Observer Bias and Moral Intuitions: Adding Fuel to Sinnott-Armstrong's Fire. *Neuroethics*, 1(2), 133–144. doi: 10.1007/s12152-008-9015-7

- Nado, J. (2014). Philosophical Expertise. *Philosophy Compass*, 9(9), 631–641.
- Nado, J. (2015). Philosophical expertise and scientific expertise. *Philosophical Psychology*, 28(7), 1026–1044.
- Nado, J. (2016, March). The intuition deniers. *Philosophical Studies*, 173(3), 781–800. doi: 10.1007/s11098-015-0519-9
- Nado, J. (2017, May). Demythologizing intuition. *Inquiry*, 60(4), 386–402. doi: 10.1080/0020174X.2016.1220639
- Nagel, J. (2012a). Gendler on Alief. *Analysis*, 72(4), 774–788.
- Nagel, J. (2012b, November). Intuitions and Experiments: A Defense of the Case Method in Epistemology. *Philosophy and Phenomenological Research*, 85(3), 495–527. doi: 10.1111/j.1933-1592.2012.00634.x
- Nagel, J., San Juan, V., & Mar, R. A. (2013). Lay Denial of Knowledge for Justified True Beliefs. *Cognition*, 129, 652–661.
- Nation, I. (2006, October). How Large a Vocabulary is Needed For Reading and Listening? *Canadian Modern Language Review*. doi: 10.3138/cmlr.63.1.59
- Nichols, S., & Pinillos, N. Á. (2018). Skepticism and the Acquisition of “Knowledge”. *Mind and Language*, 33(4), 397–414.
- Parke, C. (2018, October). *Academic journal accepts ‘Feminist Mein Kampf’* [Text.Article]. <https://www.foxnews.com/us/academic-journal-accepts-feminist-mein-kampf>.
- Perry, J. (1979). The problem of the essential indexical. *Noûs*, 3–21.
- Phipps, A. (2018, October). *What the ‘grievance studies’ hoax is really about*. <https://www.timeshighereducation.com/opinion/what-grievance-studies-hoax-really-about>.
- Pietroski, P. M. (2018). Overture. In *Conjoining Meanings*. Oxford: Oxford University Press. doi: 10.1093/oso/9780198812722.003.0001
- Plag, I. (2003). *Word-formation in English*. Cambridge; New York: Cambridge University Press.
- Plakias, A. (2019, October). Publishing without belief. *Analysis*, 79(4), 638–646. doi: 10.1093/analys/anz040
- Pluckrose, H., Lindsay, J. A., & Baghossian, P. (2018, October). *Academic Grievance Studies and the Corruption of Scholarship*. <https://areomagazine.com/2018/10/02/academic-grievance-studies-and-the-corruption-of-scholarship/>.
- Plunkett, D. (2015). Which concepts should we use?: Metalinguistic negotiations and the methodology of philosophy. *Inquiry*, 58(7-8), 828–874.
- Plunkett, D., & McPherson, T. (2019). Conceptual Ethics and The Methodology of Normative Inquiry. In Burgess, H. Cappelen, & D. Plunkett (Eds.), *Conceptual Ethics and Conceptual Engineering*. Oxford University Press.
- Potts, C. (2007, July). Into the Conventional-Implicature Dimension. *Philosophy Compass*, 2(4), 665–679. doi: 10.1111/j.1747-9991.2007.00089.x
- Pust, J. (2000). *Intuitions as evidence*. Taylor & Francis.
- Putnam, H. (1975). The meaning of ‘meaning’. *Minnesota Studies in the Philosophy of*

- Science*, 7.
- Rakedzon, T., Segev, E., Chapnik, N., Yosef, R., & Baram-Tsabari, A. (2017). Automatic jargon identifier for scientists engaging with the public and science communication educators. *PLoS ONE*, 12(8). doi: 10.1371/journal.pone.0181742
- Ramos, J. P. M., Alcock, L., Lew, K., Rago, P., Sangwin, C., & Inglis, M. (2019). Using Corpus Linguistics to Investigate Mathematical Explanation. In E. Fischer & M. Curtis (Eds.), *Methodological Advances in Experimental Philosophy* (pp. 239–263). London: Bloomsbury Academic.
- Ranalli, C. (2020). The puzzle of philosophical testimony. *European Journal of Philosophy*, 28(1), 142–163.
- Recanati, F. (2004). *Literal meaning*. Cambridge ; New York: Cambridge University Press.
- Recanati, F. (2017). Contextualism and Polysemy. *Dialectica*, 71(3), 379–397. doi: 10.1111/1746-8361.12179
- Recanati, F. (2019). Why Polysemy Supports Radical Contextualism. In G. Bella & P. Bouquet (Eds.), *Modeling and Using Context* (pp. 216–222). Cham: Springer International Publishing. doi: 10.1007/978-3-030-34974-5_18
- Reid, T. (1769). *An inquiry into the human mind, on the principles of common sense*. London : Printed for T. Cadell, (successor to A. Millar) ... and T. Longman ... London; and A. Kincaid and J. Bell, Edinburgh.
- Reuter, K. (2019). Dual character concepts. *Philosophy Compass*, 14(1), e12557. doi: 10.1111/phc3.12557
- Rini, R. A. (2015). How not to test for philosophical expertise. *Synthese*, 192(2), 431–452.
- Robinson, D. N. (1995). *An intellectual history of psychology* (3rd ed.). London: Arnold.
- Ruhl, C. (1989). *On monosemy: A study in linguistic semantics*. SUNY Press.
- Rust, J., & Schwitzgebel, E. (2013). Ethicists' and Nonethicists' Responsiveness to Student E-mails: Relationships Among Expressed Normative Attitude, Self-Described Behavior, and Empirically Observed Behavior. *Metaphilosophy*, 44(3), 350–371. doi: 10.1111/meta.12033
- Sander, T. (2019, April). Two Misconstruals of Frege's Theory of Colouring. *The Philosophical Quarterly*, 69(275), 374–392. doi: 10.1093/pq/pqy054
- Sawyer, S. (2018). VI—The Importance of Concepts. In *Proceedings of the Aristotelian Society* (Vol. 118, pp. 127–147). Oxford University Press.
- Sawyer, S. (2020). Talk and Thought. In A. Burgess, H. Cappelen, & D. Plunkett (Eds.), *Conceptual Engineering and Conceptual Ethics* (pp. 379–395). Oxford University Press.
- Schäfer, M. (2018). *The semantic transparency of English compound nouns*. Berlin: Language Science Press. doi: 10.5281/zenodo.1134595
- Schnitzler, L., Smith, S. K., Shepherd, H. L., Shaw, J., Dong, S., Carpenter, D. M., ... Dhillon, H. M. (2017, January). Communication during radiation therapy education sessions: The role of medical jargon and emotional support in clarifying patient confusion. *Patient Education and Counseling*, 100(1), 112–120. doi: 10

- .1016/j.pec.2016.08.006
- Schultz, E., Cokely, E. T., & Feltz, A. (2011). Persistent bias in expert judgments about free will and moral responsibility: A test of the Expertise Defense. *Consciousness and Cognition*, *20*(4), 1722–1731.
- Schwitzgebel, E. (2009). Do ethicists steal more books? *Philosophical Psychology*, *22*(6), 711–725.
- Schwitzgebel, E., & Cushman, F. (2012). Expertise in Moral Reasoning? Order Effects on Moral Judgment in Professional Philosophers and Non-Philosophers. *Mind and Language*, *27*(2), 135–153.
- Schwitzgebel, E., & Cushman, F. (2015). Philosophers’ Biased Judgements Persist Despite Training, Expertise and Reflection. *Cognition*, *141*(August), 127–137.
- Schwitzgebel, E., & Rust, J. (2014, June). The moral behavior of ethics professors: Relationships among self-reported behavior, expressed normative attitude, and directly observed behavior. *Philosophical Psychology*, *27*(3), 293–327. doi: 10.1080/09515089.2012.727135
- Searle, J. R. (1978). Literal Meaning. *Erkenntnis*(13), 207–224.
- Seyedsayamdost, H. (2015a, July). On gender and philosophical intuition: Failure of replication and other negative results. *Philosophical Psychology*, *28*(5), 642–673. doi: 10.1080/09515089.2014.893288
- Seyedsayamdost, H. (2015b). On Normativity and Epistemic Intuitions: Failure of Replication. *Episteme*, *12*, 95–116.
- Sharon, A. J., & Baram-Tsabari, A. (2013, January). Measuring mumbo jumbo: A preliminary quantification of the use of jargon in science communication. *Public Understanding of Science*. doi: 10.1177/0963662512469916
- Shulman, H. C., Dixon, G. N., Bullock, O. M., & Colón Amill, D. (2020, January). The Effects of Jargon on Processing Fluency, Self-Perceptions, and Scientific Engagement. *Journal of Language and Social Psychology*, 0261927X20902177. doi: 10.1177/0261927X20902177
- Sliwa, P. (2012). In defense of moral testimony. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, *158*(2), 175–195.
- Soar, D. (2018, October). Short Cuts. *London Review of Books*, 21.
- Sorensen, R. A. (1992). *Thought experiments*. New York: Oxford University Press.
- Sosa, E. (1991). *Knowledge in perspective: Selected essays in epistemology*. Cambridge [England] ; New York: Cambridge University Press.
- Sosa, E. (2007). Experimental philosophy and philosophical intuition. *Philosophical Studies*, *132*(1), 99–107.
- Spera, R., & Peña-Guzmán, D. M. (2019). The Anatomy of a Philosophical Hoax. *Metaphilosophy*, *50*(1-2), 156–174. doi: 10.1111/meta.12343
- Spicer, F. (2008). Knowledge and the heuristics of folk epistemology. In V. Hendricks (Ed.), *New Waves in Epistemology*. Palgrave Macmillan.
- Sterken, R. K. (2020). Linguistic Interventions and Transformative Communicative Disruption. In H. Cappelen, D. Plunkett, & A. Burgess (Eds.), *Conceptual Engineering and Conceptual Ethics*. Oxford University Press.

- Stich, S., & Buckwalter, W. (2011). Gender and the philosophy club. *The Philosophers' Magazine*(52), 60–65.
- Stoljar, D. (2017). *Philosophical progress: In defence of a reasonable optimism* (First edition ed.). Oxford, United Kingdom: Oxford University Press.
- Subramaniam, R., Sanjeev, R., Kuruville, S., Joy, M. T., Muralikrishnan, B., & Paul, J. (2017). Jargon: A barrier in case history taking? - A cross-sectional survey among dental students and staff. *Dental Research Journal*, 14(3), 203–208.
- Swain, S., Alexander, J., & Weinberg, J. M. (2008). The Instability of Philosophical Intuitions: Running Hot and Cold on Truetemp. *Philosophy and Phenomenological Research*, 76(1), 138–155.
- Thomson, J. J. (1976a). A defense of abortion. In J. M. Humber & R. F. Almeder (Eds.), *Biomedical ethics and the law* (pp. 39–54). Springer.
- Thomson, J. J. (1976b). Killing, Letting Die, and the Trolley Problem. *The Monist*, 59(2), 204–217.
- Thomson, J. J. (1985). The Trolley Problem. *The Yale Law Journal*, 94(6), 1395–1415. doi: 10.2307/796133
- Tobia, K., Buckwalter, W., & Stich, S. (2013). Moral Intuitions: Are Philosophers Experts? *Philosophical Psychology*, 26(5), 629–638.
- Turney, P. D., & Pantel, P. (2010). From frequency to meaning: Vector space models of semantics. *Journal of artificial intelligence research*, 37, 141–188.
- Turri, J. (2016, February). Knowledge, certainty, and assertion. *Philosophical Psychology*, 29(2), 293–299. doi: 10.1080/09515089.2015.1065314
- Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, 90(4), 293–315. doi: 10.1037/0033-295X.90.4.293
- Uhlmann, E. L., Pizarro, D. A., Tannenbaum, D., & Ditto, P. H. (2009). The motivated use of moral principles. *Judgment and Decision Making*, 4(6), 476–491.
- Valid definition and meaning.* (2020, September). <https://www.collinsdictionary.com/dictionary/english/valid>.
- van Inwagen, P. (1996). It Is Wrong, Everywhere, Always, for Anyone, to Believe Anything upon Insufficient Evidence. In J. Jordan & D. Howard-Snyder (Eds.), *Faith, Freedom and Rationality* (pp. 137–154). Savage, Maryland: Rowman and Littlefield.
- van Inwagen, P. (1997). Materialism and the Psychological-Continuity Account of Personal Identity. *Nous*, 31(s11), 305–319.
- van Inwagen, P. (2010). We're Right. They're Wrong. In R. Feldman & T. A. Warfield (Eds.), *Disagreement*. Oxford University Press. doi: 10.1093/acprof:oso/9780199226078.001.0001
- Van Cleve, J. (2006, June). Reid on the Credit of Human Testimony. In J. Lackey & E. Sosa (Eds.), *The Epistemology of Testimony*. Oxford University Press. doi: 10.1093/acprof:oso/9780199276011.001.0001
- Vicente, A., & Falkum, I. L. (2017, July). Polysemy. *Oxford Research Encyclopedia of Linguistics*. doi: 10.1093/acrefore/9780199384655.013.325
- Wanderer, J. (2013, May). Anscombe's 'Teachers'. *Journal of Philosophy of Education*,

- 47(2), 204–221. doi: 10.1111/1467-9752.12021
- Weatherson, B. (2003). What good are counterexamples? *Philosophical Studies*, 115(1), 1–31.
- Weinberg, J. M. (2007). How to challenge intuitions empirically without risking skepticism. *Midwest Studies in Philosophy*, 31(1), 318–343.
- Weinberg, J. M. (2014). Cappelen between rock and a hard place. *Philosophical Studies*, 171.
- Weinberg, J. M. (2015). Humans as instruments: Or, the inevitability of experimental philosophy. In *Experimental philosophy, rationalism, and naturalism* (pp. 179–195). Routledge.
- Weinberg, J. M., & Alexander, J. (2014). The Challenge of Sticking with Intuitions through Thick and Thin. In A. R. Booth & D. P. Rowbottom (Eds.), *Intuitions* (First edition ed., pp. 187–209). Oxford, United Kingdom: Oxford University Press.
- Weinberg, J. M., Alexander, J., Gonnerman, C., & Reuter, S. (2012, April). Restrictionism and Reflection: Challenge Deflected, or Simply Redirected? *The Monist*, 95(2), 200–222. doi: 10.5840/monist201295212
- Weinberg, J. M., Gonnerman, C., Buckner, C., & Alexander, J. (2010). Are philosophers expert intuiters? *Philosophical Psychology*, 23(3), 331–355.
- Weinberg, J. M., Nichols, S., & Stich, S. (2001). Normativity and epistemic intuitions. *Philosophical Topics*, 29(1-2), 429–460.
- Wikforss, Å. (2007, November). Semantic Externalism and Psychological Externalism: Semantic Externalism and Psychological Externalism. *Philosophy Compass*, 3(1), 158–181. doi: 10.1111/j.1747-9991.2007.00107.x
- Williamson, T. (2006). Must Do Better. In P. Greenough & M. P. Lynch (Eds.), *Truth and Realism* (pp. 278–92). Oxford University Press.
- Williamson, T. (2007). *The Philosophy of Philosophy*. Malden, MA: Blackwell Pub.
- Williamson, T. (2016, September). Abductive Philosophy. *The Philosophical Forum*, 47(3-4), 263–280. doi: 10.1111/phil.12122
- Wysocki, T. (2017, June). Arguments over Intuitions? *Review of Philosophy and Psychology*, 8(2), 477–499. doi: 10.1007/s13164-016-0301-8