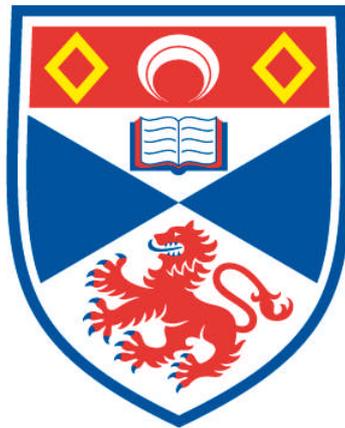


APPEALING TO INTUITIONS

Julia Langkau

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews



2013

Full metadata for this item is available in
Research@StAndrews:FullText
at:

<http://research-repository.st-andrews.ac.uk/>

Please use this identifier to cite or link to this item:

<http://hdl.handle.net/10023/3544>

This item is protected by original copyright

This item is licensed under a
Creative Commons License

Julia Langkau

Appealing to Intuitions

Thesis submitted for the degree of PhD in Philosophy
University of St Andrews, Scotland, September 2012

I, Julia Langkau, hereby certify that this thesis, which is approximately 60 000 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

I was admitted as a research student in September 2008 and as a candidate for the degree of Doctor of Philosophy in September 2008; the higher study for which this is a record was carried out in the University of St Andrews between 2008 and 2012.

6 January 2013

Signature of candidate

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

Date

Signature of supervisor

In submitting this thesis to the University of St Andrews I understand that I am 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. I also understand 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 my thesis will be electronically accessible for personal or research use unless exempt by award of an embargo as requested below, and that the library has the right to migrate my thesis into new electronic forms as required to ensure continued access to the thesis. I have obtained any third-party copyright permissions that may be required in order to allow such access and migration, or have requested the appropriate embargo below.

The following is an agreed request by candidate and supervisor regarding the electronic publication of this thesis: Access to all of printed copy but embargo of all of electronic publication of thesis for a period of 1 year on the following ground: publication would preclude future publication.

6 January 2013

Signature of supervisor

Signature of candidate

Abstract

This thesis is concerned with the ontology, epistemology, and methodology of intuitions in philosophy. It consists of an introduction, Chapter 1, and three main parts.

In the first part, Chapter 2, I defend an account of intuitions as appearance states according to which intuitions cannot be reduced to beliefs or belief-like states. I argue that an account of intuitions as appearance states can explain some crucial phenomena with respect to intuitions better than popular accounts in the current debate over the ontology of intuitions.

The second part, Chapters 3 to 5, is a reply to Timothy Williamson's (2004, 2007) view on the epistemology and methodology of intuitions. The practice of appealing to the fact that we have an intuition as evidence from thought experiments has recently been criticised by experimental philosophers. Williamson argues that since thought experiments reliably lead to knowledge of the content of our intuition, we can avoid this criticism and the resulting sceptical threat by appealing to the content of the intuition. I agree that thought experiments usually lead to knowledge of the content of our intuition. However, I show that appealing to the fact that we have an intuition is a common and useful practice. I defend the view that for methodological reasons, we ought to appeal to the fact that we have an intuition as initial evidence from thought experiments.

The third part, Chapter 6, is devoted to a paradigm method involving intuitions: the method of reflective equilibrium. Some philosophers have recently claimed that it is trivial and could even accommodate scepticism about the re-

liability of intuitions. I argue that reflective equilibrium is not compatible with such scepticism. While it is compatible with the view I defend in the second part of the thesis, more specific methodological claims have to be made.

Acknowledgements

I first and foremost wish to thank my supervisor Jessica Brown. Her careful and insightful feedback shaped this thesis significantly. For detailed comments on chapter drafts I am very grateful to Herman Cappelen, Yuri Cath, Hans-Johann Glock, Jonathan Ichikawa, Tim Mulgan, Ernest Sosa, and Brian Weatherson. I thank my examiners Katherine Hawley and Matthew McGrath, and I thank Michael De for taking the time to read and comment on the whole thesis.

During my years at the Arché Philosophical Research Centre and the St Andrews Departments of Philosophy, I was lucky to be able to work in an extremely inspiring environment. I especially thank all members and visitors of the AHRC funded ‘Intuitions and Philosophical Methodology’ project and the members of the St Andrews Friday Seminar. For being amazing friends, colleagues, and teachers, I thank Ralf Bader, Derek Ball, John Bengson, Mark Bowker, Björn Brodowski, Jessica Brown, Colin Caret, Herman Cappelen, Yuri Cath, Laura Celani, Joshua Clarkson, Paul Dimmock, Dylan Dodd, Noah Friedman-Biglin, Ephraim Glick, Patrick Greenough, Steven Hall, Katherine Hawley, Sanna Hirvonen, Andri Hjálmarsson, Ole Hjortland, Thomas Hodgson, Nick Hughes, Torfinn Huvenes, Jonathan Ichikawa, Frédérique Janssen-Lauret, Bruno Jacinto, Spencer Johnston, Dirk Kindermann, Ole Koksvik, Martin Lipman, Federico Luzzi, Toby Meadows, Dilip Ninan, Andrea Onofri, Walter Pedriali, Andrew Peet, Laura Porro, Stephen Read, Gil Sagi, Anders Schoubye, Daniele Sgaravatti, Rachel Sterken, Andreas Stokke, Margot Strohming, Paula Sweeney, Joshua Thorpe, Crispin Wright, Elia Zar-

dini, and Dan Zeman.

I also thank the Rutgers Philosophy Department which I was able to visit twice in 2011 and 2012. I thank my audiences at the University of Dublin and Trinity College Dublin, the University of Edinburgh, the University of Geneva, the Hebrew University of Jerusalem, the University of Manchester, the Western Michigan University, and the University of Zurich for helpful feedback. For financial support during the academic year 2011/2012, I am grateful to the Swiss National Science Foundation.

I thank those who gave me shelter in St Andrews several times: Laura Celani, Stephanie Cath-Garling and Yuri Cath, Derek Ball and Kat Fox, and Colin Caret. I am especially grateful to Steph and Kat for being my friends outside the philosophy world.

I thank my teachers, colleagues, and friends from Zurich: Christoph Ammann, Holger Baumann, Christoph Baumberger, Georg Brun, Hans-Johann Glock, Javier Kalhat, Dominique Kuenzle, Marion Malin, Christian Nimtz, Christoph Pfisterer, Katia Saporiti, and Peter Schulthess. I am also grateful for support and friendship from many other people who I cannot name personally.

I thank my mother Heike Turek and my grandparents Reinhilde and Klaus Turek for always supporting me. I am deeply grateful to Simone Keller and to Michael De for all their help and support and for being part of my life.

Contents

Contents	vi
1 Introduction	1
1.1 Philosophical Methodology	1
1.2 Thought Experiments	2
1.3 Intuitions	7
1.4 Intuitions as Evidence in Philosophy	10
1.5 Reflective Equilibrium	15
2 In Defence of an Inflationist Account of Intuitions	17
2.1 What are Intuitions?	17
2.2 Desiderata for an Ontology of Intuitions	19
2.3 Simple Deflationism and the Rationality Challenge	25
2.4 Elaborate Deflationism and Persistent Intuitions	30
2.5 Rationalist Inflationism and Paradigm Cases	37
2.6 Intuitions as Appearances	43
2.7 Conclusion	50
3 The Role of Intuitions in Philosophy	54
3.1 Introduction	54
3.2 Undermining the Experimentalists' Challenge?	58
3.3 Treating Intuitions as Evidence	62
3.4 Referring to Intuitions	67

<i>CONTENTS</i>	vii
3.5 Intuitions as Implicit Evidence	71
3.6 The Expertise Defence	75
3.7 Conclusion	79
4 Thought Experiments and Evidence	80
4.1 Introduction	80
4.2 Distinguishing Three Questions	83
4.3 Evidence from Thought Experiments	85
4.4 Intuitions as Initial Evidence	90
4.5 Intuitive Counterexamples	94
4.6 The Regress Argument	107
4.7 Philosophical Exceptionalism	110
4.8 Conclusion	113
5 Knowledge and Evidence in Philosophy	115
5.1 Introduction	115
5.2 Evidence in the Sciences	117
5.3 Objections and Replies	122
5.4 Evidence in Philosophy	132
5.5 Philosophical Evidence and Accessibility	143
5.6 Conclusion: Appealing to Intuitions and Scepticism	147
6 Reflective Equilibrium and Counterexamples	154
6.1 Introduction	154
6.2 The Method of Reflective Equilibrium	156
6.3 Reflective Equilibrium Applied to Counterexamples	159
6.4 Rational Belief Revision	163
6.5 Reflective Equilibrium and Scepticism	171
6.6 Reflective Equilibrium and the Appeal to Intuitions	178
6.7 Conclusion	186
Bibliography	188

Chapter 1

Introduction

1.1 Philosophical Methodology

There has recently been much discussion about the methodology of philosophy and how it compares to the methodology in other disciplines, such as the sciences. Some philosophers claim that philosophy has a distinctive methodology, namely appeal to intuitions (e.g., George Bealer [6], [7], Alvin Goldman & Joel Pust [46], Joel Pust [96], and Ernest Sosa [115]). Defenders of this methodology often take intuitions to be *sui generis* propositional attitudes that provide a genuine source of knowledge, separating philosophy from the sciences.

So-called ‘experimental philosophers’ (e.g., Weinberg, Nichols, & Stich [135], Machery, Mallon, Nichols, & Stich [81]) think that philosophers should not rely on their intuitions. They have tested laypeople’s intuitions on diverse philosophical questions in several areas of philosophical research. They found these intuitions to be sensitive to different kinds of factors irrelevant to the truth or falsity of their contents. First, they point to various types of framing and order effects people are sensitive to when presented with thought experiment scenarios.¹ Second, they have detected variations of intuitions due to eth-

¹See, e.g., Knobe [60] on how people’s judgments concerning intentionality vary depending on whether the effects of the relevant action are positive or negative, Nichols & Knobe [89] on how people’s responses to questions about moral responsibility vary depending on how the question is formulated, and Swain, Alexander, & Weinberg [119] on how some epistemic intuitions vary depending on the order in which the cases are presented.

nic differences, gender differences, socioeconomic status and other factors.² These factors are clearly irrelevant to the truth or falsity of the content of the intuitions. The evidence from experimental philosophers' studies questions the epistemic value of intuitions in general.

Some philosophers attempt to defend philosophical methodology against the attacks of experimental philosophers by denying that intuitions play an important epistemic role, or perhaps any epistemic role at all in philosophy (e.g., Timothy Williamson [138], [140], Herman Cappelen [12]). They suggest that thought experiments are best understood as involving judgements, rather than intuitions. Instead of thinking that philosophy's methodology is exceptional, they see it as continuous with methodology in the sciences. In this way, they hope to defend philosophy's methodology from a variety of attacks.

In this thesis, I look at the methodology of philosophy, and in particular defend the appeal to intuitions as evidence. I argue that intuitions are appearance states, and I defend the view that philosophers in fact do and also should appeal to intuitions as evidence in philosophy.

In what follows, I will first remind the reader of some well-known thought experiments which play a role throughout the chapters of this thesis (section 1.2). I will then give a short summary of each of the chapters. Chapter 2 is dedicated to the ontology of intuitions (section 1.3). In Chapters 3 to 5, I argue that intuitions play a role as evidence from thought experiments (section 1.4). Chapter 6 discusses the method of reflective equilibrium (section 1.5).

1.2 Thought Experiments

I will now present some thought experiments that have been used as counterexamples to philosophical theories. These thought experiments will play an important role and I present them here to refer back to them throughout the thesis. The reader familiar with the original Gettier Cases, Jackson's Mary

²For work on ethnic variations, see Machery, Olivola & de Blanc [82]; for work on gender differences, see, e.g., Buckwalter & Stich [11]; for work on the influence of the socioeconomic status, see, e.g., Weinberg, Nichols, & Stich [135].

Case, Kripke's Gödel Case, and Judith Thomson's 'Fat Man' Trolley Case might want to skip to the end of the section.

Edmund Gettier's [41] cases against the theory of knowledge as justified true belief are certainly amongst the most discussed thought experiments in philosophy. Here are both Gettier Cases as originally presented:

Gettier Case 1:

Suppose that Smith and Jones have applied for a certain job. And suppose that Smith has strong evidence for the following conjunctive proposition:

(d) Jones is the man who will get the job, and Jones has ten coins in his pocket.

Smith's evidence for (d) might be that the president of the company assured him that Jones would in the end be selected, and that he, Smith, had counted the coins in Jones's pocket ten minutes ago.

Proposition (d) entails:

(e) The man who will get the job has ten coins in his pocket.

Let us suppose that Smith sees the entailment from (d) to (e), and accepts (e) on the grounds of (d), for which he has strong evidence. In this case, Smith is clearly justified in believing that (e) is true.

But imagine, further, that unknown to Smith, he himself, not Jones, will get the job. And, also unknown to Smith, he himself has ten coins in his pocket. Proposition (e) is then true, though proposition (d), from which Smith inferred (e), is false. In our example, then, all of the following are true: (i) (e) is true, (ii) Smith believes that (e) is true, and (iii) Smith is justified in believing that (e) is true. But it is equally clear that Smith does not *know* that (e) is true; for (e) is true in virtue of the number of coins in Smith's pocket,

while Smith does not know how many coins are in Smith's pocket, and bases his belief in (e) on a count of the coins in Jones's pocket, whom he falsely believes to be the man who will get the job. [41, p. 14]

Gettier Case 2:

Let us suppose that Smith has strong evidence for the following proposition:

(f) Jones owns a Ford.

Smith's evidence might be that Jones has at all times within Smith's memory owned a car, and always a Ford, and that Jones has just offered Smith a ride while driving a Ford. Let us imagine, now, that Smith has another friend, Brown, of whose whereabouts he is totally ignorant. Smith selects three place-names quite at random, and constructs the following three propositions:

(g) Either Jones owns a Ford or Brown is in Boston.

(h) Either Jones owns a Ford or Brown is in Barcelona.

(i) Either Jones owns a Ford or Brown is in Brest-Litovsk.

Each of these propositions is entailed by (f). Imagine that Smith realizes the entailment of each of these propositions he has constructed by (f), and proceeds to accept (g), (h), and (i) on the basis of (f). Smith has correctly inferred (g), (h), and (i) from a proposition for which he has strong evidence. Smith is therefore completely justified in believing each of these three propositions. Smith, of course, has no idea where Brown is.

But imagine now that two further conditions hold. First, Jones does *not* own a Ford, but is at present driving a rented car. And second, by the sheerest coincidence, and entirely unknown to Smith, the

place mentioned in proposition (h) happens really to be the place where Brown is. If these two conditions hold then Smith does *not* know that (h) is true, even though (i) (h) *is* true, (ii) Smith does believe that (h) is true, and (iii) Smith is justified in believing that (h) is true. [41, pp. 14-15]

Another thought experiment which has provoked an extensive debate in philosophy is Frank Jackson's [54] case against physicalism. In broad terms, physicalism is the thesis that everything, including the mental, is physical. Here is how Jackson presented the Mary Case:

Mary is a brilliant scientist who is, for whatever reason, forced to investigate the world from a black and white room *via* a black and white television monitor. She specialises in the neurophysiology of vision and acquires, let us suppose, all the physical information there is to obtain about what goes on when we see ripe tomatoes, or the sky, and use terms like 'red', 'blue', and so on. She discovers, for example, just which wave-length combinations from the sky stimulate the retina, and exactly how this produces *via* the central nervous system the contraction of the vocal chords and expulsion of air from the lungs that results in the uttering of the sentence 'The sky is blue'. (It can hardly be denied that it is in principle possible to obtain all this physical information from black and white television, otherwise the Open University would *of necessity* need to use colour television.)

What will happen when Mary is released from her black and white room or is given a colour television monitor? Will she *learn* anything or not? It seems just obvious that she will learn something about the world and our visual experience of it. But then it is inescapable that her previous knowledge was incomplete. But she

had *all* the physical information. *Ergo* there is more to have than that, and Physicalism is false. [54, p. 182]

Saul Kripke [66] famously argues against the descriptive theory of the meaning of names, according to which the meaning of a name is identical to the description associated with it. Kripke gives the following case, the Gödel Case:

Imagine the following blatantly fictional situation [...] Suppose that Gödel was not in fact author of [the incompleteness] theorem. A man named ‘Schmidt’, whose body was found in Vienna under mysterious circumstances many years ago, actually did the work in question. His friend Gödel somehow got hold of the manuscript and it was thereafter attributed to Gödel. On the view in question, then, when our ordinary man uses the name ‘Gödel’, he really means to refer to Schmidt, because Schmidt is the unique person satisfying the description, ‘the man who discovered the incompleteness of arithmetic’. Of course you might try changing it to ‘the man who *published* the discovery of the incompleteness of arithmetic’. By changing the story a little further one can make even this formulation false. Anyway, most people might not even know whether the thing was published or got around by word of mouth. Let’s stick to ‘the man who discovered the incompleteness of arithmetic’. So, since the man who discovered the incompleteness of arithmetic is in fact Schmidt, we, when we talk about ‘Gödel’, are in fact always referring to Schmidt. But it seems to me that we are not. We simply are not. [66, pp. 83-84]

Judith Thomson [120], [121] presented a case which has been discussed in many different versions in moral philosophy and was originally introduced by Philippa Foot [37]. Here is the version I will refer to as the Trolley Case:

Consider a case—which I shall call Fat Man—in which you are standing on a footbridge over the trolley track. You can see a

trolley hurtling down the track, out of control. You turn around to see where the trolley is headed, and there are five workmen on the track where it exits from under the footbridge. What to do? Being an expert on trolleys, you know of one certain way to stop an out-of-control trolley: Drop a really heavy weight in its path. But where to find one? It just so happens that standing next to you on the footbridge is a fat man, a really fat man. He is leaning over the railing, watching the trolley; all you have to do is to give him a little shove, and over the railing he will go, onto the track in the path of the trolley. Would it be permissible for you to do this? Everybody to whom I have put this case says it would not be. [121, p. 1409]

Even though philosophers disagree about the nature of the judgments we make when confronted with thought experiment scenarios and whether these judgments are substantially different from other judgments or beliefs, I will call them ‘intuitions’. While Chapter 2 gives an account of what intuitions are, Chapters 3 to 6 do not rest on that account. What I argue in Chapters 3 to 6 applies to any view of what intuitions are.

1.3 Intuitions

Chapter 2, *In Defence of an Inflationist Account of Intuitions*, is concerned with the ontology of intuitions. Drawing a simplified picture of the current debate over the ontology of intuitions as propositional attitudes, there are two main camps, inflationists and deflationists. Inflationists hold that intuitions are a distinctive or *sui generis* class of mental states. Most inflationists make the further claim that the intuitions we are concerned with in philosophy provide a basic source of *a priori* knowledge (e.g., Bealer [7], Sosa [114]). I will call the latter ‘rationalist inflationists’. Deflationists think that intuitions can be reduced to judgments, beliefs, or inclinations to believe (e.g., Williamson

[138], [140]). Some deflationists argue that intuitions should not play an important epistemic role (e.g., Williamson [140], Deutsch [28]) or that they in fact do not play an epistemic role at all (Earlenbaugh & Molyneux [31]). Other deflationists think that in thought experiments (and possibly other cases), intuitions provide *a priori* knowledge (Ichikawa & Jarvis [53]). I motivate a non-rationalist inflationist account of intuitions according to which intuitions are appearance states that do not necessarily provide a source of *a priori* knowledge, because their ontology does not entail any epistemic view.

The chapter consists of three main parts. In the first part, I present some cases which display the key phenomena an ontology of intuitions should accommodate and explain. Based on these phenomena, I argue that an ontology of intuitions should:

- D1** allow for cases in which we have an intuition that P but no belief that P ;
- D2** (the *Rationality Challenge*) explain why there seems to be no failure of rationality in having inconsistent intuitions or in having an intuition that P and a belief that *not-P* (where we know of the inconsistency) whereas there would be such a failure in the case of beliefs;
- D3** explain why an intuition is often resistant to a conflicting belief.

I take it that a theory which can account for more or all desiderata should be favoured over a theory that can account for less of these desiderata. Moreover, a theory which can in some sense account better for these desiderata should be favoured over a theory which cannot account equally well for them. However, I do not take D1 to D3 to provide jointly sufficient constraints on a correct account of intuitions.

In the second part, I first look at a simple deflationist view according to which intuitions are beliefs or judgments. Understanding why the *Rationality Challenge* poses a problem to the simple deflationist will help us to evaluate Timothy Williamson's [138], [140] elaborate deflationist account according to which intuitions are either beliefs or inclinations to believe. I show that a

disambiguation of the term ‘inclination to believe’ allows for two different accounts of inclinations to believe: doxastic inclinations to believe and merely psychological inclinations to believe. The first account fits in naturally with a deflationist view according to which intuitions reduce to beliefs or belief-like states, but fails to meet D2 for similar reasons as simple deflationism. The second account meets D2, however, both the first and the second account fail to meet D3. I then argue that rationalist inflationist accounts of intuitions (e.g., Bealer [7], Sosa [115]) cannot cover all relevant cases.

In the third part, I present what I call a *non-rationalist inflationist* account of intuitions as appearance states, based on William Tolhurst’s [122] account of seemings. I show that an account of intuitions as appearance states meets D1 to D3 better than deflationist and rationalist inflationist accounts.

Let me make a couple of additional remarks concerning Chapter 2. It seems that philosophers who disagree about the ontology of intuitions disagree deeply about meta-philosophical and methodological matters. Their respective view on what intuitions are is tightly related to what they think the nature of philosophy is and to how they answer key methodological questions. Bealer [7] thinks that philosophy is autonomous in that central questions in philosophy can in principle be answered without relying substantively on the sciences. What characterizes the questions philosophers are concerned with is their universality, generality, and necessity. Intuitions play a fundamental role in philosophy in that they provide a justificatory procedure for the answers to these kinds of questions. Sosa’s [115] view on intuitions is an integral part of his general virtue epistemology, according to which our intuitions and perceptions are manifestations of our epistemic competence. For Williamson [140], the ontology of intuitions follows from his thesis that philosophy is not in principle different from the sciences or from ordinary thinking. Philosophy therefore does not need a particular justificatory procedure such as intuitions.

What ultimately decides on which account of intuitions we endorse may thus depend on many other views we have and on our general approach to

philosophy. I can obviously not do justice to arguments which derive from such other views or general approaches. I therefore focus on the question of how several accounts of intuitions can accommodate and explain certain key phenomena with respect to intuitions. Hence, I take it that, *other things being equal*, a theory which can accommodate and explain more or all phenomena or which can in some sense account better for these phenomena should be favoured over a theory that can accommodate and explain less of these phenomena or which cannot account equally well for them.

1.4 Intuitions as Evidence in Philosophy

Chapters 3 to 5 address the role of intuitions as evidence and are a reply to Williamson's [138], [140] view on evidence from thought experiments as counterexamples to philosophical theories. For the purpose of these chapters, I will presuppose the following aspects of Williamson's epistemology. First, all evidence is propositional.³ Second, all evidence is knowledge and hence factive: although we might treat false propositions as evidence sometimes, only true propositions actually are evidence.⁴

The claim that intuitions play a role as evidence is ambiguous. It can either mean that the content of the intuition or that the fact that we have an intuition is evidence. While Williamson thinks that the content of the intuition should be appealed to, I argue that we ought to use the psychological fact that we have an intuition as evidence. In what follows, I will sometimes simply use 'an intuition' for 'the fact that we have an intuition'.

The three chapters approach the topic in three steps. In Chapter 3, I defend the claim that intuitions play a role as evidence against critics who think that they do not play such a role. I argue in Chapter 4 that our practice shows that intuitions play a role as evidence and I address some worries Williamson [138], [140] raises with respect to this practice. In Chapter 5, I argue that we ought

³See Williamson [137, pp. 194-200].

⁴See Williamson [137, pp. 200-207].

to appeal to the fact that we have an intuition for methodological reasons.

I will now give a short overview of each of Chapters 3 to 5.

In Chapter 3, *The Role of Intuitions in Philosophy*, I am concerned with the role of intuitions as evidence from thought experiments. Advocates of the recent movement of experimental philosophy such as Weinberg, Nichols, & Stich [135] or Machery, Mallon, Nichols, & Stich [81] think that philosophers rely on intuitions as evidence, and they argue that for a variety of empirical reasons, this is not a good practice. According to Williamson [140], appealing to the fact that we have an intuition (rather than to the content of the intuition) as evidence from counterexamples is an unnecessary practice, and he agrees with the experimental philosophers that we ought not pursue it.

Max Deutsch [28], [29] and Joshua Earlenbaugh & Bernard Molyneux [31] go one step further and aim to undermine the experimental philosophers' arguments. Deutsch argues that in relevant and frequently discussed cases, we do not treat the fact that we have an intuition as evidence. Earlenbaugh & Molyneux defend the view that intuitions do not play any evidential role at all. If they were right, the experimental philosophers' arguments would be pointless. However, I argue that Deutsch and Earlenbaugh & Molyneux fail to undermine the experimental philosophers' challenge.

I first address Earlenbaugh & Molyneux's view. They think that whereas we trust other people's sensory states, memories, introspections, etc., we only trust our own intuitions. They give two arguments, one against the claim that we use the fact that we have an intuition as evidence and one against the claim that we use the content of the intuition as evidence. I argue that both are wrong.

I then argue against Deutsch that he uses the wrong criterion to decide whether we treat the fact that we have an intuition as evidence in philosophy. We have to distinguish between using the word 'intuition' and referring or appealing to intuitions. Even if we do not use the word 'intuition', we might

still appeal to the fact that we have an intuition as evidence. I finally show that even if we do not appeal to the fact that we have an intuition as evidence, we might still implicitly rely on it. I conclude that whether intuitions play a role as evidence can neither be read off the premises of our arguments nor decided on the basis of what we explicitly appeal to.

If the experimental philosophers' challenge cannot be undermined, we have to face it. In the last part of the chapter, I present reasons brought forward by Williamson [141] and others to think that empirical findings about the unreliability of intuitions are not relevant to philosophers' intuitions and that scepticism about the reliability of intuitions in philosophy is not warranted. I argue that since our intuitions are nevertheless fallible, psychology and experimental philosophy can sometimes help us to answer philosophical questions.

In Chapter 4, *Thought Experiments and Evidence*, I argue that appealing to the fact that we have an intuition as evidence from thought experiments is not a bad practice if this evidence is used as *initial* evidence against the target theory only.

I look at different views concerning the evidence we gain from thought experiments as counterexamples to philosophical theories. In the current debate over the methodology of thought experiments, the following three questions have not always been distinguished carefully.

- (1) What kind of facts can undermine a philosophical theory?
- (2) What kind of evidence do we gain from a thought experiment?
- (3) What kind of evidence from a thought experiment ought we to appeal to?

Concerning each of these three questions, two different answers are possible.

- (a) facts about the subject matter the theory is about
- (b) facts about our psychological states with contents about the subject matter the theory is about

Concerning (1), the standard answer is (a). There is more disagreement with respect to the answers to questions (2) and (3). I present two views that have been discussed in the current debate: a view according to which the answer to both questions is (b) (a view commonly ascribed to experimental philosophers), and an inclusive view according to which the answer to (2) is both (a) and (b) and the answer to (3) is (a) (Williamson [138], [140]). I introduce a third view, according to which the answer to (2) is both (a) and (b) and the answer to (3) is (b), i.e., according to which we ought to appeal only to the fact that we have an intuition as initial evidence against the target theory.

To show that this view corresponds with our practice, I provide examples of how we discuss intuitive counterexamples to philosophical theories some of which presuppose that we appeal to the fact that we have an intuition. For instance, we deny the relevance of certain intuitions due to some psychological disposition of the person having the intuitions, we sometimes give pragmatic or psychological explanations as to why we have an intuition, we explain why our intuitions are irrelevant with respect to the subject matter under investigation, we show that the content of our intuition is in fact not inconsistent with the target theory, and we give independent arguments in favour of the theory we aim to defend.

I then argue that two worries Williamson [138], [140] raises do not apply to this view: appealing to the fact that we have an intuition as initial evidence does not lead to a regress of giving evidence for evidence and hence does not provoke scepticism, and it does not presuppose or lead to a view according to which philosophy is different in principle from the sciences.

In chapter 5, *Knowledge and Evidence in Philosophy*, I argue that a methodological constraint on evidence in philosophy can explain why we appeal to the fact that we have an intuition as evidence from thought experiments.

I first present two cases, *Sally the Teacher* and *Sam the Psychologist*, which suggest that there are methodological constraints on evidence in the sciences.

In these cases, Sally and Sam know a proposition that P but are not in a good enough epistemic position to assert that P in a scientific context. The reason seems to be that they have only non-scientific evidence, i.e., evidence not generated through recognized scientific methods. In other words, Sally and Sam do not have the right kind of evidence.

I then argue that philosophy is similar to the sciences with respect to some relevant aspects, which suggests that there are methodological constraints on evidence in philosophy as well, and I discuss what these constraints could be.

Williamson [138], [140] suggests two ways to explain why contemporary philosophers appeal to intuitions on the basis of certain properties of this evidence. In *The Philosophy of Philosophy* [140], he argues that many sceptical contemporary philosophers appeal to the fact that they have an intuition as evidence because they falsely believe in *Evidence Neutrality*, which is the idea that a community of philosophers can always in principle achieve common knowledge as to whether any given proposition constitutes evidence or not. In ‘Philosophical ‘Intuitions’ and Scepticism About Judgement’ [138], Williamson mentions why even some non-sceptical philosophers might want to appeal to the fact that we have an intuition instead of to the content of the intuition: they believe evidence has to meet an operational standard, and they think that the fact that someone has an intuition can meet this standard. I reject both *Evidence Neutrality* and an operational standard as methodological norms.

I then suggest a norm for evidence in philosophy according to which we ought to appeal to the part of our total evidence for a certain claim which is relatively easy to access. This norm can explain why we appeal to the fact that we have an intuition as evidence from a thought experiment even if we are not sceptical about knowledge from thought experiments.

Chapters 3 to 5 are a reply to Williamson’s view on evidence from thought experiments in the following sense. According to Williamson, referring to intuitions as evidence is a practice we should not pursue, because it opens a gap

between facts about our psychological states and the truth of the contents of these states, which provokes scepticism. Since, as I argue, we appeal to the fact that we have an intuition for methodological reasons only, it follows that we have to close the gap for methodological reasons only. This does not affect our knowledge, and hence does not constitute a sceptical threat. While we have knowledge of a proposition which undermines the target theory, we might nevertheless not have succeeded in refuting the theory. However, the history of philosophy shows that it is indeed very hard to refute a philosophical theory.

1.5 Reflective Equilibrium

Reflective equilibrium has been considered a paradigm philosophical method involving intuitions. It has been extensively discussed in normative ethics and political philosophy. The key idea of reflective equilibrium as introduced by John Rawls [99], [100] for moral and political philosophy is that we test our moral judgments (or intuitions) and moral principles (or theories) against each other and revise and refine both when they are inconsistent. In the literature, ‘reflective equilibrium’ ambiguously stands for the *method* of reflective equilibrium and for the *state* in which all our beliefs are in equilibrium, i.e., coherent. I am concerned with the method of reflective equilibrium.

Without specifying what exactly they mean by it, philosophers in all areas of research frequently use the term ‘reflective equilibrium’ when they mention the methods and aims of their inquiry. It has been suggested that ‘reflective equilibrium’ is nothing more than a metaphor for the rational performance of philosophy: for taking into account all relevant information available and for working out the most plausible, coherent, and comprehensive theory of the subject matter under investigation. Michael DePaul [25] and Peter Singer [104] suggest that it could even be compatible with the kind of scepticism about the reliability of intuitions advocated by some experimental philosophers, i.e., with a view according to which we ought not take any intuitions into account at all.

I am interested in the question whether reflective equilibrium is as trivial as some philosophers think or whether it could give us useful methodological advice. If it gives us methodological advice, having followed reflective equilibrium could provide us with some evidence as to the truth of the resultant theory which is relatively easy to access and hence is evidence we ought to appeal to (as I argue in Chapter 5).

The chapter consists of two main parts. In the first part, I present reflective equilibrium as it has been discussed in moral and political philosophy and apply it to epistemology, more precisely to a fictional philosopher Sophie who thinks about the Gettier Cases. I then specify what philosophers mean when they say that reflective equilibrium is trivial.

In the second part, I look at different views philosophers hold in the current debate concerning the role of intuitions in philosophy. I argue that the sceptical view according to which intuitions ought not play any role in philosophy is not compatible with reflective equilibrium. I moreover argue that the view I defend in Chapters 3 to 5, according to which we ought to appeal to the fact that we have an intuition as initial evidence from thought experiments, is compatible with reflective equilibrium.

I conclude that MRE is not as trivial as some philosophers think, but more specific methodological claims have to be made in order for it to serve as a methodological guide.

Chapter 2

In Defence of an Inflationist Account of Intuitions

Abstract

This chapter is concerned with the ontology of intuitions. I first develop desiderata for an ontology of intuitions on the basis of paradigm cases. I then discuss whether some popular accounts of intuitions meet these desiderata: the view that intuitions are simply beliefs, Timothy Williamson's [138], [140] account according to which intuitions are either beliefs or inclinations to believe, and George Bealer's [7] and Ernest Sosa's [115] rationalist accounts of intuitions. I finally present a theory of intuitions as appearance states, based on William Tolhurst's [122] account of seemings, which I argue meets the suggested desiderata best.

2.1 What are Intuitions?

In reflecting on their methodology, philosophers use the terms 'intuition' and 'seeming' in various ways. Some use them to refer to a source of knowledge in general, some use them for a source of *a priori* knowledge only. Others use the terms to refer to a kind of propositional attitude, some of which use them for common sense judgments. Psychologists sometimes speak of intuition

when they theorize about an unconscious, gut-guided decision process (e.g., Gigerenzer [42]). There are even more ways of using ‘it is intuitive’ and ‘it seems’ in ordinary language, some of which occur in philosophy papers, such as the use as a hedge or a cautious way of presenting evidence.¹

In this chapter, I am not concerned with different uses or meanings of ‘intuition’, ‘intuitive’, ‘seeming’, ‘seems’, or similar terms. I am interested in *the kind of propositional attitude* we are dealing with in some specific contexts, such as thought experiments as counterexamples to philosophical theories. We say we have the intuition that a person in a Gettier Case against the theory of knowledge as justified true belief (the JTB theory) has the justified true belief that p but no knowledge that p .² In Frank Jackson’s Mary Case against physicalism, we say we have the intuition that colour scientist Mary learns something when she leaves her black and white room and sees something coloured for the first time in her life.³ Other paradigm cases in which we use the term ‘intuition’ in philosophy are the lottery paradoxes, the sorites paradoxes, or the paradox of emotional response to fiction⁴, or other paradox cases. We say we have an intuition towards each proposition of the inconsistent set of propositions that constitutes a paradox. I will restrict my considerations to the use of ‘intuition’ or related terms in thought experiments and paradoxes, and I will use ‘ S has an intuition that P ’ and ‘ P is intuitive to S ’ interchangeably for these cases.

Drawing a simplified picture of the current debate over the ontology of intuitions as propositional attitudes, there are two main camps, inflationists and deflationists. Inflationists hold that intuitions are a distinctive or *sui generis* class of mental states. Most inflationists make the further claim that the intuitions we are concerned with in philosophy provide a basic source of *a priori* knowledge (e.g., Bealer [7], Sosa [114]). I will call the latter ‘rationalist inflationists’. Deflationists think that intuitions can be reduced to judgments,

¹See Cappelen [12, pp. 25-93].

²Gettier [41].

³Jackson [54]. Most philosophers think that the Mary Case is not a genuine counterexample to physicalism, including Jackson [56].

⁴I will return to these paradoxes below.

beliefs, or inclinations to believe (e.g., Williamson [138], [140]⁵). Some deflationists argue that intuitions should not play an important epistemic role (e.g., Williamson [140], Deutsch [28]) or that they in fact do not play an epistemic role at all (Earlenbaugh & Molyneux [31]). (I address these arguments and discuss the role of intuitions as evidence in Chapters 3, 4, and 5.) Other deflationists think that in thought experiments (and possibly other cases), intuitions provide *a priori* knowledge (Ichikawa & Jarvis [53]). I motivate a non-rationalist inflationist account of intuitions according to which intuitions are appearance states that do not necessarily provide a source of *a priori* knowledge. Their ontology does not entail any epistemic view. This option has been widely ignored in the current debate over the ontology and epistemic role of intuitions.

The chapter consists of three main parts. In the first part (section 2.2), I present 6 cases which show some key phenomena an ontology of intuitions should accommodate and explain. A special focus lies on cases that have been subject to extensive first-order philosophical debate but have received little attention in the current debate over the ontology of intuitions. In the second part (sections 2.3-2.5), I discuss a simple deflationist view according to which intuitions are just beliefs or judgments (section 2.3), Timothy Williamson's [138], [140] more elaborate deflationist view according to which intuitions are either beliefs or inclinations to believe (section 2.4), and George Bealer's [7] and Ernest Sosa's [115] rationalist inflationist accounts of intuitions (section 2.5). In the third part of the chapter (section 2.6), I present a non-rationalist inflationist account of intuitions as appearance states which I show can accommodate and explain the relevant phenomena presented in the first part best.

2.2 Desiderata for an Ontology of Intuitions

Let me give some examples of paradigm cases in which philosophers use the term 'intuition'. Any presentation of these cases will be disputable to some extent. In order to avoid begging the question against deviant descriptions, let

⁵Williamson attributes a similar view to Lewis [74] and van Inwagen [126].

me introduce a hypothetical philosopher Helen. Helen's apprehension of the cases might not be shared by everyone. However, philosophers in the current debate about intuitions have described themselves as being in similar situations with respect to these cases or with respect to similar cases. In this sense, Helen is a representative contemporary philosopher.

At first sight, some of the cases look similar because they are ones in which Helen has an intuition whose content she does not believe. However, there are differences which will play a role in my arguments below, and it will become clear in the course of the chapter why we need all these cases.

Here are three well-known thought experiments and Helen's reaction to them.

Case 1: Gettier Case

Smith has applied for a job. He has a justified belief that someone else, Jones, will get the job, and he also has a justified belief that Jones has 10 coins in his pocket. Smith therefore justifiably believes that the man who will get the job has 10 coins in his pocket. In fact, Jones does not get the job and Smith himself does. As it happens, Smith unknowingly also has 10 coins in his pocket. His belief that the man who will get the job has 10 coins in his pocket therefore is justified and true. Does Smith also know that the man who will get the job has 10 coins in his pocket?⁶

Helen *has the intuition* that Smith does not know that the man who will get the job has 10 coins in his pocket. It is also true that she *believes* that Smith does not know that the man who will get the job has 10 coins in his pocket.

Case 2: Trolley Case

A trolley is hurtling down a track towards five people. You are on a bridge under which it will pass, and you can stop it by dropping

⁶Gettier [41].

a heavy weight from the bridge. There is a very fat man standing next to you on the bridge, and the only way to stop the trolley is to push him over the bridge and onto the track, killing him in order to save the five people on the track. Should you push the fat man over the bridge?⁷

Helen *has the intuition* that, was she in the situation described, she should not push the fat man over the bridge. However, she knows that the numerous Trolley Cases are subject to extensive debates in moral philosophy, and not knowing these cases and debates in much detail, she feels like she cannot decide what to believe. She therefore refrains from judging. Helen has the intuition but she *does not believe* that she should not push the fat man over the bridge.

Case 3: Mary Case

Mary is a scientist who is forced to investigate the world from a black and white room via a black and white television monitor. She specializes in the neurophysiology of vision and has all the physical information there is to obtain about what goes on when we see colours. Will Mary learn anything new when she is released from her black and white room and sees colours for the first time?⁸

Helen *has the intuition* that Mary will learn something new when released from her black and white room. However, she *firmly believes* that physicalism is true, and she believes that it is inconsistent with the truth of physicalism that Mary will learn something new.

Let us now look at three paradox cases.

Case 4: Lottery Paradox

p *S* knows that *S* will not have enough money to go on a safari this year.

⁷Thomson [120], [121].

⁸Jackson [54].

q If S knows that S will not have enough money to go on a safari this year, then S is in a position to know that S will not win a major prize in a lottery this year.

r S is not in a position to know that S will not win a major prize in a lottery this year.⁹

All three propositions of the Lottery Paradox considered individually *are intuitive* to Helen, but she notices that they are jointly inconsistent. She cannot decide what to believe and *refrains from judging*, just as in Case 2.

Case 5: Paradox of the Heap

p A pile of sufficiently many grains is a heap.

q One single grain cannot make a difference to whether something is a heap or not.

r A single grain is not a heap.

All three propositions of the Paradox of the Heap considered individually *are intuitive* to Helen, but she notices that they are jointly inconsistent. Now, suppose Helen thinks she has a solution to the Paradox of the Heap, which leads her to believe that q is false, i.e., she *believes* that it is false that one single grain cannot make a difference to whether something is a heap or not.¹⁰ It is still intuitive to Helen that one single grain cannot make a difference to whether something is a heap or not.

Case 6: Paradox of Emotional Response to Fiction

p We can feel genuine pity for the fictional character Anna Karenina.

⁹Hawthorne [51, p. 2].

¹⁰See, e.g., Williamson [136].

q We cannot feel pity for someone we believe does not exist.

r We believe (even when engaged with fiction) that the relevant fictional characters do not exist.

All three propositions of the Paradox of Emotional Response to Fiction considered individually *are intuitive* to Helen, but she notices that they are jointly inconsistent. Suppose that Helen endorses a theory of quasi-emotion (similar to Walton's [129]) and *rejects p* on the basis of it: we cannot feel *genuine* pity for Anna Karenina. It is, however, still intuitive to Helen that *p*, i.e., it is still intuitive to her that we can feel genuine pity for Anna Karenina.¹¹

Let me now point to some phenomena with respect to Cases 1 to 6. In Case 1, Helen has the intuition that Smith in the Gettier Case does not have knowledge, but it is also true that she believes that Smith does not have knowledge (most contemporary epistemologists share this belief with Helen). This might lead one to the view that her intuition ought to be identified with her belief. In the contemporary debate over the ontology and epistemic role of intuitions, Gettier Cases have been referred to as paradigm cases of intuition driven philosophy.¹² Taking them to be paradigmatic might incline one to think that intuitions in general are just beliefs or belief-like states. However, there are some more phenomena that can be observed with respect to our cases.

¹¹See, e.g., Radford [97], Gendler & Kovakovich [38]. It is pretty much common sense that *r* should not be rejected. It does not seem plausible to assume that when we read a novel such as Tolstoy's *Anna Karenina*, we temporarily cease to believe that Anna Karenina does not exist. Radford, e.g., rejects this sort of solution to the paradox. With respect to Shakespeare's *Romeo and Juliet*, he says: 'Of course we don't ever forget that Mercutio is only a character in a play, but we "suspend our disbelief" in his reality. The theatre management and the producer connive at this. They dim the lights and try to find good actors. They, and we, frown on other members of the audience who draw attention to themselves and distract us by coughing, and if, during a scene, say a stage hand steals on [sic.], picks up a chair that should have been removed and sheepishly departs, our response is destroyed. The "illusion" is shattered. All this is true but the paradox remains. When we watch a play we do not direct our thoughts to it's only being a play. We don't continually remind ourselves of this—unless we are trying to reduce the effect of the work on us. Nonetheless [...] we are never unaware that we are watching a play, and one about fictional characters even at the most exciting and moving moments. So the paradox is not solved by invoking "suspension of disbelief", though it occurs and is connived at.' [97, pp. 71-72].

¹²See, e.g., Williamson [140].

First, Cases 2 to 6 show that we do not necessarily believe what is intuitive to us. Second, they show that it does not seem irrational for Helen to have an intuition that P without believing that P .

With respect to this second phenomenon, the cases are slightly different from each other. In Case 2, Helen has an intuition and suspends belief. In Case 4, in which Helen has not solved the paradox, believing p , q , and r altogether would seem irrational, because p , q , and r form a contradictory set of propositions. However, finding each of the propositions intuitive seems not irrational. In Cases 3, 5, and 6, in which Helen has a firm belief that contradicts her intuition, having a belief that P and a belief that $not-P$ at the same time would seem clearly irrational. However, finding P intuitive and at the same time believing $not-P$ seems not irrational.

An ontology of intuitions should explain why it seems that Helen is not irrational in these cases. I will call the challenge to explain this phenomenon the *Rationality Challenge* to an ontology of intuitions. The *Rationality Challenge* has not received much attention in the current debate, for different reasons. The reason deflationists have not paid much attention to it is most likely because the Gettier Cases have been the main focus.¹³ Rationalist inflationists, in contrast, have focused on the following third phenomenon. In Cases 3, 5, and 6, Helen's intuitions are resistant to conflicting beliefs—or at least more persistent than beliefs in philosophy usually are.¹⁴ In paradoxes such as Cases 5 and 6, we usually begin finding each of a set of contradictory propositions intuitive and then search for a theory which will preserve as many of our intuitions as possible. Even if one or more of our intuitions are inconsistent with the theory we come to be convinced of, we often keep finding each proposition of the set intuitive. A similar persistence can be observed in Case 3, the Mary Case. Even though Helen is convinced that physicalism is true, she still has

¹³See Williamson [140, chapter 7].

¹⁴See, e.g., Bealer: '[N]early any proposition about which you have beliefs, authority, cajoling, intimidation, and so forth can, fairly readily, insinuate at least some doubt and thereby diminish to some extent, perhaps only briefly, the strength of your belief. But seldom, if ever, do these things so readily diminish the strength of your intuitions.' [7, p. 208].

the intuition that Mary learns something new when she considers the case.

An ontology of intuitions should account for all and not just some of these paradigm cases. Based on the distinctive features Cases 1 to 6 show, I will now suggest desiderata for an ontology of intuitions. An ontology of intuitions should:

- D1** allow for cases in which we have an intuition that P but no belief that P ;
- D2** (the *Rationality Challenge*) explain why there seems to be no failure of rationality in having inconsistent intuitions or in having an intuition that P and a belief that *not- P* (where we know of the inconsistency) whereas there would be such a failure in the case of beliefs;
- D3** explain why an intuition is often resistant to a conflicting belief.

These desiderata will provide a useful guide for the evaluation of different accounts of intuitions. I take it that, other things being equal, a theory which can account for more or all desiderata should be favoured over a theory that can account for less of these desiderata. Moreover, a theory which can in some sense account better for these desiderata should be favoured over a theory which cannot account equally well for them. However, I do not take D1 to D3 to provide jointly sufficient constraints on a correct account of intuitions.

In what follows, I argue that deflationist and rationalist inflationist accounts of intuitions do not meet D1 to D3 as well as a *non-rationalist* inflationist account of intuitions which I present in section 2.6.

2.3 Simple Deflationism and the Rationality Challenge

According to simple deflationism, intuitions are just judgments or beliefs. Simple deflationism is not popular in the current debate over intuitions in philosophy, and it is easy to see why deflationists hold more elaborate views. I will address D1 and D3 only briefly and then focus on D2, the *Rationality Challenge*. Understanding why the *Rationality Challenge* poses a problem to the

simple deflationist will help us to evaluate an elaborate deflationist account in section 2.4.

It is obvious that simple deflationism cannot meet D1, the desideratum that an account of intuitions allow for cases in which we have an intuition that P but no belief that P . In fact, D1 begs the question against the simple deflationist, because she disagrees with our description of the cases in section 2.2. The simple deflationist would have to say that while it might seem that in the Trolley Case (Case 2), Helen suspends belief, she actually does believe that we should not push the fat man, and in Cases 3 to 6 (the Mary Case and the paradox cases), Helen simply has conflicting beliefs. However, I take it that the description I gave of the above cases is *prima facie* plausible. At the very least, the burden of proof rests on the simple deflationist to explain why Cases 2 to 6 are ones where we believe that P while it clearly seems that we have no belief that P .

But even if we grant the simple deflationist a description of the cases in terms of beliefs only, she owes us an explanation of why intuitions are often resistant to conflicting beliefs (D3) and why there seems to be no failure of rationality in Cases 3 to 6 (D2). With respect to D3, she could argue that resistance to conflicting beliefs is not special to the beliefs we call ‘intuitions’, but is a common psychological phenomenon with respect to belief in general. Ideally, the simple deflationist could give an explanation of why our beliefs are sometimes resistant to conflicting beliefs and sometimes not. Other things being equal, a theory which can provide such an explanation is to be favoured over a theory which cannot. (As we will see in sections 2.4 and 2.5, inflationist accounts can give such an explanation.)

I will now focus on the *Rationality Challenge* (D2), which says that there seems to be no failure of rationality in Cases 3, 5, and 6. I will show that two obvious strategies to deal with this phenomenon within a simple deflationist view of intuitions are problematic. First, the simple deflationist could identify Helen’s intuitions with degrees of belief: in Cases 2 to 6, Helen does not fully

believe the propositions, she rather believes them to a certain degree only, and a certain distribution of credences makes her beliefs rational.¹⁵

However, there are good reasons to reject an account of intuitions in terms of degrees of belief. According to an account of degrees of belief, a subject S is not rational if the degrees of her contradictory beliefs P and $not-P$ do not add up to 1 on a scale from 0 to 1. *Prima facie*, such an account of intuitions cannot accommodate Cases 3, 5, and 6. The reason is that in these cases, our philosopher Helen is fully convinced of some proposition $not-P$ while she also has an intuition that P . Given that Helen is fully convinced that $not-P$, her credence in $not-P$ is close to 1. If she is rational, then her credence in P will be close to 0, which cannot explain that she has an intuition that P . Phenomenologically, an intuition is just not like a belief with very low credence. If, however, Helen's credence in P is not close to 0, then she does not assign a coherent distribution of credences to her beliefs, in which case she is irrational, thus failing to meet the *Rationality Challenge*.

Second, in order to meet the *Rationality Challenge*, the simple deflationist could reject the *Conjunction Principle*, which says that if we are rational, then if we believe a set of propositions, we also believe the conjunction of them. Rejecting this principle, the simple deflationist could hold that whereas believing a conjunction of inconsistent propositions is irrational, believing each proposition individually is not irrational. Here is the principle:

Conjunction Principle: If S is rational, then if S believes P and S believes Q , then S believes (P and Q).

If intuitions are beliefs, the *Conjunction Principle* implies that belief in some contradictions is rational—paradoxes being paradigm cases. Consider Helen. She is rational in cases 3, 5, and 6, but by the conjunction principle, she believes in a contradiction. However, it is widely accepted that any theory of rational belief is in need of a principle that prohibits belief in a contradiction¹⁶:

¹⁵This view derives from what Sturgeon [118] calls the *Fine View* of rational belief.

¹⁶Evnine [33, p. 202].

No Contradiction Principle: If S is rational, then S does not believe (P and $not-P$).

Obviously, the *No Contradiction Principle* is in conflict with the *Conjunction Principle*, so we have to abandon one of them. Rejecting the *No Contradiction Principle* is a possible way to go, but it is certainly the least popular way to go.¹⁷ In the literature on rational belief, some philosophers discard the *Conjunction Principle*.¹⁸ This strategy allows us to rationally believe all propositions of a paradox individually without believing the conjunction of them. The simple deflationist could make use of this strategy in order to explain why we do not seem irrational in Cases 3 to 6.

Rejecting the *Conjunction Principle* commits one to the claim that there is a difference between being in a state of having contradictory beliefs and being in a state of believing a contradiction, so that one could individually believe P and Q without thereby believing the conjunction (P and Q). Simon Evnine [33] argues that there is evidence to think that believing (P and Q) and believing P and believing Q amounts to being in the same state. Evnine observes that it is presupposed by our practice of attributing conjunctive beliefs that there is only one state. For instance, when we summarize a person's view, we do this by attributing a conjunctive belief. We do not require evidence that some psychological process has occurred in which all the individual beliefs were conjoined to one single belief. According to Evnine, this suggests that it is not possible to believe the conjuncts without believing the conjunction.¹⁹ If it is not even possible to believe the conjuncts without believing the conjunction, then the rejection of the *Conjunction Principle* fails.²⁰

It seems, however, plausible that a subject S can, without being irrational, hold beliefs that are *implicitly* contradictory, i.e., beliefs that imply P and *not-*

¹⁷For a rejection of the *No Contradiction Principle*, see Priest [93].

¹⁸Kyburg [68], Cherniak [15], Kornblith [63], Foley [36], and more recently Christensen [16] and Sturgeon [118].

¹⁹Evnine takes it to be a conceptual truth that believing the conjunctions and believing the conjunction is one and the same state, see Evnine [33, p. 215].

²⁰Evnine [33, pp. 214-219].

P without S knowing or being aware of it. We often hold several beliefs which are jointly inconsistent either for a lack of knowledge or a lack of reflection, and it is at least controversial whether we are irrational in these cases. The cases under consideration, in contrast, are cases of conscious belief in which we are aware of the inconsistency in our set of beliefs. That we are perfectly aware of our intuitions towards all three propositions of the paradox cases and the fact that they are inconsistent is obvious from the fact that we consider Cases 4, 5, and 6 as paradoxical. The same is true for Case 3, the Mary Case: the intuition, the belief, and the fact that they are inconsistent are conscious to our philosopher Helen. All we need for our cases is a principle that prohibits conscious belief in a conscious contradiction (P and *not*- P). Hence, even if general arguments against it fail, the *Conjunction Principle* should still apply to cases of a small number of conscious beliefs where we are also aware of the inferential relations between these beliefs—cases such as paradoxes and thought experiments.

I have argued that two *prima facie* plausible strategies to defend simple deflationism against the *Rationality Challenge* (D2) are problematic. Neither of the strategies provide a satisfying explanation of why we are not irrational in cases where we have contradictory or conflicting intuitions. The deflationists in the current debate over the ontology of intuitions seem to acknowledge that a more elaborate view is required. In the following section, I present and discuss Williamson's account, according to which an intuition is either a belief or an inclination to believe. I consider that a disambiguation of the term 'inclination to believe' allows for two different accounts. The account that fits in naturally with a deflationist view fails to meet D2 for reasons similar to the ones discussed with respect to simple deflationism. Whereas the second account meets D2, inflationist accounts (which I discuss in sections 2.5 and 2.6) have an advantage over it: they provide an explanation of why intuitions sometimes are resistant to conflicting beliefs, i.e., they meet D3.

2.4 Elaborate Deflationism and Persistent Intuitions

According to Williamson [138], [140], to have the intuition that P is either to believe that P or to be merely consciously inclined to judge or believe that P . I will call this view an ‘elaborate deflationist’ view of intuitions. Elaborate deflationism applies to our cases as follows. In Case 1, the Gettier Case, Helen believes that Smith does not know that the man who will get the job has 10 coins in his pocket; in Cases 2 to 6, she is merely consciously inclined to believe the respective propositions.

Unlike simple deflationism, Williamson’s elaborate deflationism allows for cases in which we have an intuition but no belief, i.e., it meets D1. At first sight, it also provides an explanation of why we are not irrational in Cases 3 to 6, i.e., it seems to meet the *Rationality Challenge* (D2). According to Williamson, if S believes that P , S is committed to the truth of P , but if S is merely consciously inclined to believe that P , S is not committed to the truth of P :

In the way in which I am committed to the propositions that I believe, I am not committed to the propositions that I am merely inclined to believe; I am merely inclined to commit myself to them in that way. [138, p. 128]

Because there is no commitment to the truth involved in inclinations to believe, they do not obey any *No Contradiction Principle*. In Case 4, the Lottery Paradox, we have conscious inclinations to believe all three propositions of the inconsistent set. Since we are not committed to the truth of these propositions, the inconsistency does not pose a problem. In Cases 3, 5, and 6, we have a conscious inclination to believe that P and a belief that $not-P$. The explanation of why we are not irrational in these cases is again that we are not committed to the truth of the content of our conscious inclinations to believe that P . In what follows, I will first argue that elaborate deflationism cannot meet the *Rationality Challenge* because it cannot account for the fact

that we are equally rational in both kinds of cases (Cases 3, 5, and 6 on the one hand, and Case 4 on the other hand). Second, I will argue that elaborate deflationism does not meet D3.

To argue that elaborate deflationism cannot meet the *Rationality Challenge*, let us see if a conscious inclination to believe is necessary and sufficient for an intuition. There are not many arguments to be found in the literature as to whether our intuitions necessarily involve conscious inclinations to believe or not. Williamson makes the following introspective claim about intuitions:

I can feel such an inclination even if it is quite stably overridden, and I am not in the least danger of giving way to temptation (just as one can feel the inclination to kick someone without being in the least danger of giving way). [140, p. 217]

Unlike Williamson, some authors think that we sometimes have an intuition that P without a conscious inclination to believe that P . Earlenbaugh & Molyneux [31] defend an account of intuitions as inclinations to believe, but they think that it seems plausible that there are cases of intuitions without inclinations to believe:

There at least appear to be cases where one has an intuition without an accompanying inclination to believe. It seems coherent, that is, to claim that we can have an intuition that p without being inclined to believe p . [31, p. 106]

As an example of such a case, they mention the *Naive Comprehension Axiom* (i.e., for any property, there is a set of things having that property) which is intuitive to us even though we have a firm belief that it is not true. This case is parallel to our Cases 3, 5, and 6, where Helen has a firm belief that *not-P* and an intuition that P . I share Earlenbaugh & Molyneux's view that there appear to be cases where one has an intuition that P and no inclination to believe that P , and I think that there in fact are such cases.

The reason why philosophers disagree about whether there are intuitions without inclinations to believe could be due to an ambiguity of the term ‘inclination to believe’. Let me introduce a distinction between ‘doxastic inclinations to believe’ and ‘merely psychological inclinations to believe’. Doxastic inclinations to believe work in certain respects like guesses. If S guesses that Q , S is not committed to the truth of Q . However, if S firmly believes *not- Q* and S is rational, she does not guess that Q . Her guess that Q commits her to stop guessing that Q as soon as she firmly believes *not- Q* . Similarly, if S firmly believes *not- Q* and S is rational, she is not doxastically inclined to believe that Q .²¹ Doxastic inclinations to believe are derivatives of belief in the sense that they involve some kind of rational commitment, i.e., they inherit some rationality principles from belief. They certainly do not inherit the *No Contradiction Principle*, since the fact that we do not violate this principle when we have contradictory conscious inclinations to believe is exactly the work inclinations to believe do as opposed to beliefs. However, doxastic inclinations to believe are likely to inherit *Closure under Consciously Known Entailment*. For belief, the principle is formulated thus:

Closure under Consciously Known Entailment (for conscious belief): If S is rational, then if S consciously believes that P and knows that Q is entailed by P , and considers whether Q , then S consciously believes that Q .

Closure under Consciously Known Entailment is not affected by the work inclinations to believe do as opposed to beliefs. Suppose that conscious belief is closed under consciously known entailment. For instance, suppose that Helen consciously believes that if the cat is in the kitchen, it is not in the garden. If Helen also consciously believes that the cat is in the kitchen, then if she is rational, Helen consciously believes that it is not in the garden. For a doxastic inclination to believe, the principle is formulated thus:

²¹There are, of course, differences between doxastic inclinations to believe and guesses. Whereas we can have contradictory doxastic inclinations to believe, we cannot guess that P and guess that *not- P* .

Closure under Consciously Known Entailment (for conscious doxastic inclination to believe): If S is rational, then if S is consciously doxastically inclined to believe that P and consciously knows that Q is entailed by P , and considers whether Q , then S is consciously doxastically inclined to believe that Q .

Suppose that Helen wonders where the cat is and for some reason comes to be consciously inclined to believe that the cat is in the kitchen, but she does not believe it (say, because her evidence is rather poor). Suppose further that Helen now considers whether the cat is in the garden. Then if she is rational, she will be consciously inclined to believe that the cat is not in the garden. This case makes it at least plausible that *Closure under Consciously Known Entailment* holds not only for conscious beliefs but *also* for conscious doxastic inclinations to believe. Merely psychological inclinations to believe, in contrast, do not involve any kind of rational commitment: they are not subject to rational evaluation, i.e., we might want to call them arational.

We could resolve the disagreement between Williamson and Earlenbaugh & Molyneux as follows. What Williamson claims is that he can feel a merely psychological inclination to believe that P when he has an intuition that P and a belief that *not-P*. Earlenbaugh & Molyneux's, however, have the intuition that we do not necessarily have a doxastic inclination to believe that P when we have an intuition that P and a belief that *not-P*.

I will now show that a doxastic inclination to believe that P is neither sufficient nor necessary for an intuition that P in the relevant cases where an intuition is not a belief. To begin with, there is at least a *prima facie* difference between cases where we have an intuition that P and a firm belief that *not-P* on the one hand (Cases 3, 5, and 6) and cases where we have an intuition that P without a firm belief that *not-P* on the other hand (Case 4). If we think that intuitions are doxastic inclinations to believe, we should expect it to seem at least more irrational to be consciously inclined to believe P in the case in which we have the firm belief that *not-P* than in the case in which we only

have a contradicting conscious inclination to believe *not-P*. After all, we are committed to the truth of our firm belief in Cases 3, 5, and 6, whereas we are not committed to the truth of our inconsistent conscious inclinations in Case 4. However, all cases are similar with respect to rationality: we seem not irrational in Cases 3, 5, and 6 to the same extent as in Case 4.

To show that a doxastic inclination to believe is not sufficient for an intuition, let us now have a look at Case 4 again. Suppose we are doxastically inclined to believe the contents of our intuitions in the Lottery Paradox. Here are the propositions again:

p: *S* knows that *S* will not have enough money to go on a safari this year.

q: If *S* knows that *S* will not have enough money to go on a safari this year, then *S* is in a position to know that *S* will not win a major prize in a lottery this year.

r: *S* is in not in a position to know that *S* will not win a major prize in a lottery this year.

Helen has intuitions towards *p*, *q*, and *r*. Suppose intuitions are doxastic inclinations to believe. Then Helen has doxastic inclinations to believe *p*, *q*, and *r*. It is also plausible to suppose that Helen, who is a rational subject and knows what a paradox is, consciously knows that, for instance, (*p* and *r*) entails *not-q*. Hence, Helen has a doxastic inclination to believe *q* and a doxastic inclination to believe *not-q*, but clearly she has no intuition that *not-q*, which shows that a doxastic inclination to believe is not sufficient for an intuition.

In order to show that a doxastic inclination to believe is not *necessary* for an intuition either, let us have a look at cases in which Helen has an intuition that *P* and a firm belief that *not-P* (Cases 3, 5, and 6). In these cases, Helen does not believe the content of her intuition that *P*. If she had a doxastic inclination to believe that *P*, she would moreover have a doxastic inclination to

believe what she knows to be entailed by P . However, Helen is not doxastically inclined to believe what is entailed by the content of her intuition that P . Take Case 3 as an example. Helen has the firm belief that physicalism is true and the intuition that Mary will learn something new. However, Helen clearly does not have the intuition that physicalism is false, and she does not even have the doxastic inclination to believe that physicalism is false. After all, Helen has a firm belief that physicalism is true. But doxastic inclinations to believe, I argued above, obey *Closure under Consciously Known Entailment*. So, if Helen has a doxastic inclination to believe that Mary learns something new, and she knows that this entails that physicalism is false, then Helen should have the doxastic inclination to believe that physicalism is false. Hence, since the state we are talking about does not obey *Closure under Consciously Known Entailment*, it does not seem to involve a doxastic inclination to believe. Helen only has an intuition that Mary will learn something new, and this intuition does not obey *Closure under Consciously Known Entailment*.

If what I argued is correct, it supports Earlenbaugh & Molyneux's intuition: there are cases where we have an intuition that P and no inclination to believe that P , at least no *doxastic* inclination to believe. One might, however, want to argue that elaborate deflationism does not involve *doxastic* inclinations to believe but only *merely psychological* inclinations to believe. The fact that these inclinations to believe are arational explains why we are not irrational in the relevant cases.

In order to defend this version of the elaborate deflationist account of intuitions, one would have to say more about two things. First, one would have to explain in what sense such an account of intuitions as beliefs or merely psychological inclinations to believe would be a *deflationist* account, given that merely psychological inclinations to believe do not seem to share many of belief's characteristic properties. Second, one would have to say something about what exactly merely psychological inclinations to believe are, given that they do not share much with belief.

Now let us come back to the *Rationality Challenge*. An account of intuitions as beliefs or doxastic inclinations to believe simply does not cover all our cases. In Cases 3, 5, and 6, we do not have a doxastic inclination to believe, and this explains why we are not irrational. However, a version of elaborate deflationism according to which intuitions are beliefs or merely psychological inclinations to believe would meet the *Rationality Challenge*. It would explain why we are equally rational in Cases 3, 5, and 6 on the one hand and Case 4 on the other hand.

But even if the elaborate deflationist can establish that intuitions are either beliefs or merely psychological inclinations to believe, it seems that she cannot provide an explanation of why intuitions are often resistant to conflicting beliefs. This would mean that elaborate deflationism cannot meet D3.

Here is, however, one explanation a deflationist might offer. There might be a simple psychological explanation for merely psychological inclinations to believe: sometimes we are psychologically inclined to believe a proposition P even if we are convinced that P is false. For instance, a mother is inclined to believe that her son is innocent of killing the neighbour's cat even if she firmly believes that he is guilty, maybe even if she saw him killing the cat. This merely psychological inclination to believe is probably caused by how much she loves her son and by her wish that he had not killed the cat. Since her love and her wish persist, her inclination is persistent as well. Maybe a similar explanation could work for some of our intuitions in philosophy? This seems rather implausible. First, we usually do not have philosophical intuitions towards personal subject matters we are emotionally involved in (except maybe in moral philosophy). Second, we share many intuitions with other philosophers, and intuitions such as the ones we have in paradox cases are equally persistent to many philosophers. It is very implausible to assume that we would find a psychological explanation for all philosophers with regard to these cases. Moreover, and as we will see in section 2.5, the inflationist has a better explanation.

Note that the phenomenon of persistent intuitions concerns both thought experiments (as in Case 3) and paradoxes (as in Cases 5 and 6). It does not occur randomly, it rather appears systematically with paradigm cases of our use of the term ‘intuition’ in philosophy. The fact that intuitions are resistant to conflicting beliefs has been one of the main arguments for inflationist accounts of intuitions. In the next two sections, I present and address two inflationist accounts.

2.5 Rationalist Inflationism and Paradigm Cases

George Bealer’s (e.g., [6], [7]) and Ernest Sosa’s (e.g., [115]) inflationist accounts of intuitions tie in with the rationalist tradition after Descartes according to which intuition is understood as a source of *a priori* knowledge.²² Whereas traditional rationalism understands this source as infallible, contemporary rationalists like Bealer and Sosa think that our intuitions sometimes lead us awry. They can thus account for the fact that we have an intuition towards each of the propositions that form a paradox, but at least one intuition is misguided.

According to Bealer, intuitions constitute a ‘*sui generis*, irreducible, natural [...] propositional attitude that occurs episodically’.²³ When we have an intuition that *P*, it *intellectually seems* to us that it must be that *P* and could not be otherwise. This phenomenology cannot be analyzed further, which Alvin Plantinga expresses as follows:

[Intuitions come with] that peculiar form of phenomenology with which we are all well acquainted, but which I can’t describe in any way other than as the phenomenology that goes with seeing that such a proposition is true. [92, pp. 105-106]

At first sight, Bealer’s rationalist inflationist account meets desiderata D1 to D3. In fact, the phenomena behind D1, D2, and D3 have motivated rationalist

²²See also Plantinga [92].

²³Bealer [7, p. 207].

inflationist views that intuitions are an extra kind of propositional attitude.²⁴ If intuitions are a *sui generis* propositional attitude, this explains that there are cases in which we have an intuition that P but no belief that P (D1), that we are not irrational in the case of a set of inconsistent intuitions or an intuition that P and a belief that *not-P* (D2), and that intuitions are more persistent than beliefs (D3). Bealer points to the fact that similar phenomena can be observed with respect to sensory seemings in optical illusions such as the Müller-Lyer illusion. In the Müller-Lyer illusion, it seems to us that one line is longer, but we do not believe it. We are not irrational when it seems to us that one line is longer, and the illusion is persistent.²⁵

However, looking at our cases, it seems that in philosophy, we are not only concerned with what would qualify as an intellectual seeming. Compare Cases 5 and 6, the Paradox of the Heap and the Paradox of Emotional Response to Fiction. Whereas the rationalist inflationist would take it that the intuitions in Case 5 are a source of *a priori* knowledge, she would certainly agree that the propositions in Case 6 are contingent and that they cannot be known *a priori*, and hence that we cannot have rational intuitions towards them.

Bealer, however, distinguishes two kinds of intuitions, intellectual or rational *a priori* seemings and non-rational *a posteriori* seemings. He calls non-rational seemings ‘physical’ because they stem from the experience we have of the contingent physical world. Their content is not presented as necessary, and in philosophy, we are not concerned with them:

When we speak here of intuition, we mean “rational intuition.”
 This is distinguished from what physicists call “physical intuition.”
 We have a physical intuition that, when a house is undermined, it will fall. This does not count as a rational intuition, for it does not present itself as necessary: it does not seem that a house undermined *must* fall; plainly it is *possible* for a house undermined to

²⁴Bealer [7, p. 208], [6, p. 6], Sosa [112, p. 258f], Pust [96, p. 32f].

²⁵See, e.g., Bealer [7, pp. 207-214].

remain in its original position or, indeed, to rise up.’ [7, p. 207]

Other cases where we have non-rational intuitions are scientific thought experiments. Bealer mentions the following case:

A classic example is Newton’s thought experiment about a rotating bucket in an otherwise empty space. Would water creep up the side of the bucket (assuming that the physical laws remained unchanged)? Rational intuition is silent about this sort of question. Rational intuitions concern such matters as whether a case is possible (logically or metaphysically), and about whether a concept applies to such cases. [7, p. 207]

Another example is Schroedinger’s case of the cat which is directed against the Copenhagen interpretation of quantum mechanics. Schroedinger asks us to imagine a cat, placed in a sealed box together with some poison, a radioactive source and a Geiger counter. If the Geiger counter detects radiation, the poison gets released. The Copenhagen interpretation of quantum mechanics implies that after a while, the cat is simultaneously dead and alive. However, we have the intuition that when we look into the box, the cat is either dead or alive, but it cannot be both. Our intuition in this case has contingent content we get to know *a posteriori*.²⁶

Bealer thinks that intuitions we are dealing with in philosophy are different from intuitions in the sciences: they are rational seemings.²⁷ However, while our intuitions in the Paradox of Emotional Response to Fiction would not be considered rational seemings, this case is certainly relevant to philosophy. Since its introduction by Colin Radford [97], it has played a central role in the philosophy of fiction and the philosophy of emotions. Some philosophers argue that *q*, the claim that we cannot feel pity for someone we do not believe exists, is false (e.g., Carroll [13], Feagin [34]). Other philosophers deny *p*, the claim that we can feel pity for Anna Karenina. Some of them think that we cannot have

²⁶See Sorensen [110] for more thought experiments in the sciences.

²⁷Bealer [7, p. 207].

a *genuine* emotional response to fictional situations or fictional characters, and that what we experience is a quasi-emotion (e.g., Walton [130]). Others think that the emotions we experience when we engage with fiction are really directed towards relevantly similar actual characters and situations (e.g., Charlton [14]). Obviously, the Paradox of Emotional Response to Fiction poses a problem we are discussing in philosophy.²⁸ It seems that non-rational intuitions not only play a role in scientific thought experiments, they also play a role in philosophy.

In order to justify an ontological difference between our intuitions in Case 5 and our intuitions in Case 6, Bealer could argue that the phenomenology of our intuitions in Case 5 is different from the phenomenology of our intuitions in Case 6. On introspection, I cannot find any difference between my intuition that a single grain cannot make a difference to whether something is a heap or not on the one hand and my intuition that we can feel genuine pity for Anna Karenina on the other hand. Even philosophers who share Bealer's epistemic view disagree about the supposed phenomenological aspect of rational intuitions. Joel Pust [96], who endorses a rationalist inflationist view very similar to Bealer's, thinks that the content of a rational intuition only seems necessary to us *on reflection*.²⁹ However, even on reflection, I do not experience such a difference.

Sosa's [115] account of rational intuitions does not presuppose a particular phenomenology. Sosa is mainly interested in the question of how our intuitions in paradigm cases of *a priori* knowledge—such as in arithmetic, geometry, or in philosophical thought experiments and paradoxes—justify our beliefs, and how we can best account for these intuitions. According to Sosa, intuitions are *intellectual seemings* that, unlike perceptual seemings, are not passively received and hence are epistemically evaluable.³⁰ A visual experience as if

²⁸A similar case where we have intuitions with contingent, *a posteriori* content is the Paradox of Painful Art which consists of the following three propositions: 1. People do not seek out situations that arouse painful emotions. 2. People have painful emotions in response to some art. 3. People seek out art that they know will arouse painful emotions. See Smuts [108, p. 60].

²⁹Pust [96, pp. 31-39].

³⁰Sosa [115, pp. 46, 49].

there is a fire, for instance,

[...] is thought to yield foundational justification, being itself beyond relevant evaluation, beyond justification and unjustification. The intellectual seeming, by contrast *is* thus evaluable. A reason can be assigned the wrong weight, as it attracts one's assent too much, or too little. [115, p. 49]

Sosa specifies conditions under which an intellectual seeming is *intuitive* and *rational*.³¹ These conditions have to be understood in the context of his broader epistemic view, i.e., his virtue epistemology:

[...] the proposed account [of intuitions] has two parts: first, an understanding of intuitions as a special sort of intellectual seemings, *intuitive* seemings; second, a definition of the sort of intuition that is distinctively "rational."

1. An intellectual seeming is *intuitive* when it is an attraction to assent triggered simply by considering a proposition consciously with understanding. (Of course, one may so much as understand the proposition only through a complex and prolonged process that includes perception, memory, testimony, or inference.)
2. S rationally intuits that p if and only if S's intuitive attraction to assent to < p > is explained by a competence (an epistemic ability or virtue) on the part of S to discriminate, among contents that he understands well enough, the true from the false, in some subfield of the modally strong (the necessarily true or necessarily false), with no reliance on introspection, perception, memory, testimony, or inference (nor further reliance,

³¹Note that I am concerned with Sosa's characterization of intuitions in his 2007 book *A Virtue Epistemology* (Vol. 1) only, see [115, pp. 44-63].

anyhow, than any required for so much as understanding the given proposition). [115, pp. 60-61]

One might think that since Sosa characterizes intuitions as attractions to assent that are epistemically evaluable, his view is similar to an elaborate deflationist account according to which intuitions are beliefs or doxastic inclinations to believe. In this case, it would not meet D2 and D3 for the reasons given in section 2.4. If intuitions are epistemically evaluable, it seems that we cannot explain why someone with a firm belief that *not-P* and an intuition that *P* is rational (D2) and why intuitions are more resistant to contradictory beliefs than other beliefs (D3). However, it seems that Sosa rather characterizes intuitions primarily as intellectual seemings, and then specifies them as *intuitive* seemings that are a certain kind of attraction and that share with beliefs the property of being rationally evaluable.

Just like Bealer, Sosa is only interested in rational *a priori* intuitions and how they justify our beliefs. My question, however, is what could explain the phenomena behind D1 to D3 best. Assuming that rational intuitions do not or at least do not necessarily come with a particular phenomenology, we can say the following. The intuitions we have towards the propositions in Case 6 do not seem different from the intuitions we have in Case 5 with respect to D1 to D3. Independent of whether intuitions provide a source of *a priori* or of *a posteriori* knowledge, whether their content is modally strong or not, or whether they ought to be explained by a certain competence, the following holds. First, intuitions do not necessarily involve a belief with the same content (D1). Second, we are not irrational when we have an intuition that *P* and a belief that *not-P* (D2). Third, intuitions are more persistent than beliefs (D3).

It seems that the ontology of our intuitions in Case 5 should be the same as the ontology of our intuitions in Case 6. I will now show how an account of intuitions as appearances can meet D1 to D3 best, because it covers all our cases. The advantage of this non-rationalist inflationist account is that it does not commit us to any particular epistemic view.

2.6 Intuitions as Appearances

In a short paper titled ‘Seemings’, William Tolhurst [122] gives a very dense presentation of his general account of seemings:

Philosophers have worried much about whether, when, and why we are justified in believing that things are as they seem. Less attention has been paid to the general question of what it is for things to seem to be a certain way. Most work on the nature of seemings has focussed on perception and, to a lesser extent, memory [...]. Some seemings do not fit into these two categories. For example, a student in a logic class may consider a proof and find it seems to be valid. Seemings whose objects are abstract objects are not covered by accounts of memory and perception. In what follows, I seek to develop a general account that will cover all these cases. [122, p. 293]

While I cannot discuss every aspect of Tolhurst’s account of seemings, I will look at how it could apply to Cases 1 to 6 and meet desiderata D1 to D3.

According to Tolhurst, seemings are very much like beliefs in that they are intentional states with a mind-to-world direction of fit (unlike desires, wishes, etc.). Their contents either fit or fail to fit how things are, just like belief contents. Despite their similarity to beliefs, seemings constitute a distinct kind of mental state. Their relation to beliefs is both causal and epistemic: a seeming can cause and also evidentially support a belief. The circumstances under which a seeming both causes and evidentially supports a belief, however, do not always obtain. Whether or not they obtain, the seeming involves a felt demand that one believe its content: a felt veridicality, ‘the feel of a state whose content reveals how things really are.’³² Seemings also carry with them a disposition to form second order beliefs about them.

Here is Tolhurst’s characterization of seemings:

³²Tolhurst [122, pp. 298-299].

It seems to S that ϕ is t only if

- i) S is in a mental state, m , whose content is that ϕ is t,
- ii) m has a property (felt veridicality) which
 - a) evokes in S the felt demand that S believe that ϕ is t,
 - and
 - b) would, other things equal, evoke in S the felt demand that S believe that m is veridical if S were to reflect on her experience of m . [122, p. 299]

Let us first look at some cases Tolhurst gives to distinguish seemings from beliefs and to show how seemings cause and evidentially support beliefs. Instead of ‘felt demand’, I will use ‘felt inclination’ in what follows (Tolhurst uses both terms interchangeably. I take a ‘felt inclination’ to be the same as a ‘conscious inclination’, see my discussion of Williamson’s account in section 2.4).

Case A: *Seeming with belief*

It seems to Sam that there is a puddle in front of her and Sam also believes that there is a puddle in front of her. The seeming both causes and evidentially supports her belief.³³

Case B: *Seeming with felt inclination to believe*

Bert has conflicting seemings. On the one hand, Ernie looks angry. On the other hand, Ernie assures Bert that he is not angry. It seems (*prima facie*) to Bert that Ernie is angry, but it also seems (*prima facie*) to him that he is not angry. The first inclines Bert to believe that Ernie is angry, the second inclines Bert to believe that Ernie is not angry. Because of his conflicting felt inclinations, Bert withholds judgment. Both seemings are overridden and generate no beliefs.³⁴

³³Tolhurst [122, p. 294].

³⁴Tolhurst [122, pp. 294-295].

Case C: *No belief, no felt inclination to believe: mere appearance*

William, who is fully aware of the situation, is looking at a white object under red light. It appears to him that the object is red, but this appearance is overridden by his knowledge to the contrary, and he is not inclined to believe that the object is red. Whereas the appearance disposes William to believe that the object is red, there is no psychological motivation for him to believe that the object is red.³⁵

Tolhurst distinguishes between all-things-considered seemings which generate beliefs, as in Case A, and *prima facie* seemings which only generate felt inclinations to believe, as in Case B. Case C cannot be explained by Tolhurst's characterization of seemings given above because the appearance that the object is red does not involve a felt inclination to believe that the object is red. To cover cases like C, Tolhurst introduces the term 'appearance' for states regardless of whether they incline us to believe that things are as they appear. He then uses the term 'seeming' for appearances that cause an inclination to believe or that generate a belief and the term 'mere appearance' for those that do not cause an inclination to believe and that do not generate a belief. Tolhurst thinks that ordinary usage does not provide an answer to whether states that incline us to believe the content and such that do not should be classed together or not. However, it seems obvious that he in fact does class them together. Tolhurst gives the following characterization of appearances in general.³⁶

All appearances dispose the subject to believe the content [...] This is not to say that they all provide occurrent motivation, e.g., a felt

³⁵Tolhurst [122, pp. 295-297].

³⁶One reason why Tolhurst does not introduce seemings as kinds of appearance states from the beginning might be that his main interest are *prima facie* seemings which carry with them a felt inclination to believe. *Prima facie* seemings are more interesting than mere appearances and more salient than seemings all-things-considered. Mere appearances are less interesting because they are overridden and hence of no epistemic relevance. Seemings all-things-considered are less salient because they are accompanied by a belief and there is less reason to address the seeming if there is a belief with the same content.

inclination, to believe; but that they create conditions favourable for believing. This disposition, like other dispositions, is only activated under appropriate circumstances. Things that are brittle have a disposition to crack and shatter: under some circumstances they crack, under more extreme conditions they shatter, but often neither happens because these conditions do not obtain. Likewise, under appropriate circumstances appearances generate beliefs and felt inclinations to believe. When an appearance is decisively defeated or passes unnoticed, the conditions appropriate for the activation of the disposition do not obtain and it may not provide actual psychological support for believing. [122, p. 296]³⁷

The analogy to brittle things that sometimes crack and sometimes shatter suggests that under some circumstances, an appearance generates a felt inclination to believe, in which case it is a *prima facie* seeming, and under some other circumstances, it generates a belief, in which case it is a seeming all-things-considered. Under some circumstances, however, it does not generate either an inclination to believe or a belief, in which case we call it a ‘mere appearance’.

In line with Tolhurst’s characterization of a seeming, we can give the following more general characterization of an appearance state.

It appears to S that p only if

- i) S is in a mental state, m , whose content is that p ,
- ii) m has a property (felt veridicality) which

³⁷Gregory [49] gives a similar account of the relation between sensory appearances and beliefs: ‘There are especially intimate links between sensory appearances and many of our beliefs about the outside world. Indeed, the links are so intimate that some philosophers have identified sensory appearances with beliefs, or with the acquisition of beliefs. That identification is too strong – the appearances persist when weve not got any inclination to trust them – but there is something right about it: part of what it is for sensory appearances to be ‘appearances’ is that they can be accurate or inaccurate in just the way that beliefs can be.’ Gregory [49, p. 321]. Gregory then takes sensory and memory appearances as a starting point to argue that possibility appearances are distinctive states as well.

- a) would, absent a contradicting belief that *not-p*, evoke in *S* the felt inclination that *S* believe that *p*, and
- b) would, absent a contradicting felt inclination to believe (and hence absent a belief³⁸) that *not-p*, evoke in *S* the belief that *p*, and
- c) would, other things equal, evoke in *S* the felt demand that *S* believe that *m* is veridical if *S* were to reflect on her experience of *m*.

Tolhurst also distinguishes ‘peripheral appearances’, which are mere appearances that pass unnoticed by the subject but may be registered in memory at later times. Note that peripheral appearances do not satisfy the characterization I have given above, because they are not occurrent or felt states. I am not sure that peripheral appearances should be classed together with felt appearances, because it is not clear what it means that something appears to a subject to be a certain way without the subject actually noticing that it seems to her in a certain way.

Note again that Tolhurst’s seemings can have all kinds of contents, and so can appearances in general. However, all appearances have propositional content, they involve a felt veridicality, and they dispose us to have a felt inclination to believe and to have a belief with the same content. These conditions are necessary, but they might not be sufficient to specify all kinds of appearances. Sensory appearances, for instance, might come with a special phenomenology. More work has to be done to distinguish different kinds of appearances, but this would go beyond the scope of the present chapter.

Having presented and slightly adapted Tolhurst’s account of appearances, let us now see whether, as an account of intuitions, it can meet D1 to D3. Note that we are not interested in the justificatory relation our beliefs might have to

³⁸It is widely held and I take it to be the case that whenever a subject *S* has a belief that *P* she also has an inclination to believe that *P*.

the appearance states. We are merely interested in the causal relation which might explain the phenomena observed with respect to Cases 1 to 6.

Case 1, the Gettier Case, is parallel to Case A. It appears to Helen that Smith does not know that the man who will get the job has 10 coins in his pocket. This seeming causes her belief that the man who will get the job has 10 coins in his pocket. The circumstances are such that nothing blocks Helen's belief.

Case 4, the Lottery Paradox, is not directly parallel to any of the cases, but it is closest to Case B. Helen has appearances towards a set of inconsistent propositions, not just towards two contradictory propositions as in Case B. Since the propositions are inconsistent and she realizes that they cannot possibly all be true, Helen refrains from judging. However, nothing prevents her from being inclined to believe each proposition. Having an inconsistent set of appearances could have the same effect as having two contradictory appearances: it blocks belief, but not inclination to believe.

Case 2, the Trolley Case, is not directly covered by any of Tolhurst's cases either. It is again most similar to Case B. It appears to Helen that if she was in the situation described in the case, she should not push the fat man. We could describe the case such that Helen is inclined to believe that she should not push the fat man, or we could describe the case such that Helen is not even inclined to believe that she should not push the fat man. What blocks Helen's belief or even her inclination to believe is that she finds the question difficult to answer, and she thinks that in order to form a belief, she would have to think a lot more about moral theory. Hence, a lack of confidence in our judgment in a certain area or the knowledge that we have not fully understood a complex matter might block our belief that P or even our inclination to believe that P .

Cases 3, 5, and 6 are parallel to Case C. The Mary Case, the Paradox of the Heap, and the Paradox of Emotional Response to Fiction are such that Helen does not have either a felt inclination or a belief with the content of her respective intuition. In Case 3, it appears to Helen that Mary will learn

something new when released from her black and white room. However, this appearance does not cause a belief or an inclination to believe. The reason is that it is blocked by her firm belief that physicalism is true and her belief that physicalism is inconsistent with the proposition that Mary learns something new.

In Case 5, it appears to Helen that one single grain cannot make a difference to whether something is a heap or not. However, this appearance does not cause a belief or an inclination to believe. The reason is that it is blocked by Helen's firm belief that she has solved the paradox because it is false that one single grain cannot make such a difference.

In Case 6, it appears to Helen that we can feel genuine pity for Anna Karenina. However, this appearance does not cause a belief or an inclination to believe because it is blocked by Helen's firm belief that we can only have quasi-emotions towards Anna Karenina. In all three cases, Helen has a belief which contradicts the content of her appearance, and hence a belief or a felt inclination to believe with the content of the appearance does not occur.

Let us come back to desiderata D1 to D3 for an account of intuitions. We have seen that an account of intuitions as appearances allows for cases in which we have an intuition that P but no belief that P (Tolhurst's cases B and C and Cases 2 to 6), which means that it meets D1. However, appearances share some features with beliefs: they are intentional states with a mind-to-word direction of fit, the latter of which distinguishes them from states like wondering, assuming, supposing, and imagining. Due to the felt veridicality they involve, appearances dispose us to form a belief or a felt inclination to believe towards their content. Note that the 'inclination to believe' could either be a merely psychological inclination to believe or a doxastic inclination to believe which involves some rational commitment (as distinguished in section 2.4). It could be the latter because in cases where we have an intuition that P and a firm belief that $\text{not-}P$, we have no inclination to believe that P , we merely have an appearance as to P . Appearances do not involve any kind of rational or

doxastic commitment to their content. This explains why it is not irrational to have an intuition that P and a belief that $not-P$ at the same time. Hence, an account of intuitions as appearances can meet the *Rationality Challenge* (D2). D3 can be met as well, since the resistance of intuitions to conflicting beliefs could just be a feature of appearances in general. We are familiar with this phenomenon from the resistance of appearances in the case of optical illusions. Even though we know that the stick in the water is not bent, it appears to us that it is bent whenever we look at it.

2.7 Conclusion

I have presented and discussed several accounts of intuitions which are popular in the current debate over the ontology of intuitions, and I have argued that we have reasons to endorse an account that has not received much attention in the current debate, namely an account of intuitions as appearance states.

After having presented some key phenomena that, other things being equal, an ontology of intuitions should accommodate and explain, I discussed a simple deflationist view according to which intuitions are just beliefs or judgments, a more elaborate deflationist view according to which intuitions are either beliefs or inclinations to believe, and two rationalist inflationist accounts of intuitions. I argued that each of these accounts has some problems accommodating and explaining these key phenomena.

I showed in section 2.3 that the main problem the simple deflationist account has is that it cannot meet the *Rationality Challenge* (D2), which is the desideratum that an account of intuitions explain why it is not irrational to have an inconsistent set of intuitions or an intuition that P and a belief that $not-P$ at the same time.

In section 2.4, I showed that an elaborate deflationist view can be understood in two ways. First, intuitions are either beliefs or doxastic inclinations to believe. Second, intuitions are either beliefs or merely psychological inclina-

tions to believe. I argued that the first version of elaborate deflationism still cannot meet the *Rationality Challenge*. Moreover, both versions of elaborate deflationism cannot meet D3, the desideratum that an account of intuitions explain that our intuitions are more resistant to conflicting beliefs than beliefs in philosophy usually are.

I argued in section 2.5 that Bealer's and Sosa's rationalist inflationist accounts of intuitions cannot account for all relevant cases, and I presented a non-rationalist inflationist view in section 2.6. This account of intuitions as appearances covers all relevant cases and hence meets desiderata D1 to D3 better than rationalist inflationist accounts.

One might want to object to this account of intuitions for other reasons. For instance, if one thinks that intuitions come with no phenomenology at all, one might be unsatisfied with the fact that appearances are supposed to come with an unanalyzable phenomenology. Tolhurst says the following:

The notion of felt veridicality resists analysis. But we are acquainted with it; reflection on suitable examples should be sufficient to call it to mind. [122, p. 299]

One might deny that intuitions come with something like felt veridicality. Many philosophers in the recent debate claim not to be acquainted with any phenomenology of intuitions. Herman Cappelen [12], for instance, reports the following:

[...] I cannot, even with the best of will, discern a special feeling that accompanies my contemplation of the naive comprehension axiom, Gettier cases and other alleged paradigms of the intuitive. [12, p. 117]

Similarly, Williamson [140] denies that there is any kind of phenomenology that comes along with intuitions:

For myself, I am aware of no intellectual seeming beyond my conscious inclination to believe the Gettier proposition. Similarly, I am aware of no intellectual seeming beyond my conscious inclination to believe Naive Comprehension, which I resist because I know better.

[140, p. 217]

Williamson's comment is directed against intellectual seemings, but he would probably say the same with respect to appearances. To be consistent, however, one ought to deny that one is aware of appearances in general, not just appearances relevant in philosophy, because appearances in philosophy are not supposed to be different from any other appearances.

There is not much one can say in reply to these introspective reports, since there is no way to convince anyone that they are familiar to the notion of felt veridicality. However, all I intended to do in this chapter was to show that an account of intuitions as appearances can meet our desiderata D1 to D3. These desiderata rely on phenomena which seem to be widely shared amongst philosophers.

A further advantage of the account of intuitions I have defended over rationalist inflationist accounts is that it allows for one single state in all our cases, independent of epistemic aspects. Just as there is one belief state for beliefs with all kinds of contents, there is one state for appearances with all kinds of contents. Rationalist inflationist accounts of intuitions are only available if there is a clear-cut *a priori*–*a posteriori* distinction. It is, however, controversial whether there is a clear-cut *a priori*–*a posteriori* distinction. Moreover, it is controversial whether intuitions from thought experiments provide *a priori* knowledge even if there is such a distinction. For instance, as a consequence of Williamson's account of thought experiments, our intuitions in thought experiments have contingent contents that can only be known *a posteriori*.³⁹

³⁹See Williamson [140]. Ichikawa and Jarvis [53] take this consequence as a reason to reject Williamson's analysis of thought experiments, because the standard view has it that intuitions like the Gettier intuition are known *a priori*. However, Williamson rejects the idea that there is a clear-cut *a priori*–*a posteriori* distinction on independent grounds, so an argument that presupposes *a priori* knowledge begs the question against his analysis of

The deflationist account I have defended can explain our talk of ‘intuitive judgments’ or ‘intuitive beliefs’. These judgments or beliefs are based (causally or epistemically) on intuitions as appearance states. When we speak about ‘intuitions’, this might either involve a belief (or even knowledge) as in Case 1, or an inclination to believe as in Cases 2 and 4, or no inclination to believe at all, as in Cases 3, 5, and 6.

Chapter 3

The Role of Intuitions in Philosophy

Abstract

The practice of appealing to intuitions as evidence has recently been criticised by experimental philosophers. While some traditional philosophers defend intuitions as a trustworthy source of evidence, others try to undermine the challenge. They argue that in frequently discussed cases intuitions do not play a role as evidence (Deutsch [28], [29]), or that intuitions do not play any role as evidence in philosophy (Earlenbaugh & Molyneux [31]). I argue that these attempts to undermine the experimentalists' challenge fail. If what I argue is correct, philosophers have to face the experimental philosophers' challenge.

3.1 Introduction

Some philosophers think that intuitions are treated as evidence in contemporary analytic philosophy as well as in the history of philosophy, going back to Plato. In the recent debate about philosophical methodology, George Bealer [7], Alvin Goldman & Joel Pust [46], and Ernest Sosa [115], among others, provide several arguments for the claim that intuitions are justifiably treated so,

since they in fact are evidence in philosophy. Advocates of the current movement of experimental philosophy such as Weinberg, Nichols, & Stich [135] or Machery, Mallon, Nichols, & Stich [81], to mention only two well-known papers, agree that we rely on intuitions, but they consider it a practice which is not conducive to the aims of traditional philosophy. Weinberg, Nichols, & Stich's empirical studies on Gettier Cases show that there are significant variations in people's intuitive responses to these cases, depending on their culture and socioeconomical status. Based on this empirical result, the authors argue that traditional epistemologists do not succeed in establishing genuine epistemic norms because what they really explore are norms local to their own cultural and socioeconomic group only. Machery, Mallon, Nichols & Stich's empirical studies on Kripke's Gödel Case reveal significant variations in people's intuitions about the reference of names. Raising similar concerns as Weinberg, Nichols, & Stich, the authors draw the following conclusion:

[...] our data indicate that philosophers must radically revise their methodology. Since the intuitions philosophers pronounce from their armchairs are likely to be a product of their own culture and their academic training, in order to determine the implicit theories that underlie the use of names across cultures, philosophers need to get out of their armchairs. [81, B9]

In this chapter, I do not discuss the experimentalists' studies or their arguments against the use of intuitions as evidence in philosophy, and I do not go into much details about possible ways to defend the claim that intuitions are evidence. Rather, I discuss some traditional philosophers' reactions to the experimentalists' challenge that take, more or less, the same line of argument. Instead of defending intuitions as a source of evidence, they aim to undermine the challenge by denying that intuitions play the assigned role. Max Deutsch [28], [29] argues that in relevant and frequently discussed cases, intuitions do not play the evidential role that has been assigned to them. Joshua Earlen-

baugh & Bernard Molyneux [31] defend the view that contrary to what some traditional philosophers as well as experimental philosophers claim, intuitions have not been treated as evidence in philosophy at all. Before presenting their arguments and responding to them in detail, I will make some helpful distinctions.

Philosophers who think that intuitions play a role as evidence usually also think that thought experiments as counterexamples or potential counterexamples to philosophical theories are paradigm cases of the use of intuitions as evidence. In Frank Jackson's Mary Case against physicalism, some of us have the intuition that colour scientist Mary learns something new when she leaves her black and white room and sees something coloured for the first time in her life (Jackson [54]). In Kripke's counterexample to a descriptivist theory of the meaning of names, the Gödel Case, most of us have the intuition that 'Gödel' refers to Gödel and not to Schmidt (Kripke [66]). In the Gettier Cases against the theory of knowledge as justified true belief (the JTB theory), we have the intuition that a person in a particular situation lacks knowledge that P despite having a justified true belief that P (Gettier [41]). The defendants of the claim that intuitions play a role as evidence take intuitions to be evidence against a philosophical theory in these and similar cases.

The claim that intuitions are evidence is ambiguous, as has been stated in the literature (e.g., Lycan [79], Earlenbaugh & Molyneux [31]). It can either mean that the content of the intuition (also called 'the intuited'), or that the psychological state (also called 'the intuiting') is evidence. Philosophers who think that evidence is factive (e.g., Williamson [140]) refer to the first by speaking about a *fact about the world* as evidence, and to the second by speaking about the *fact that we have an intuition* as evidence. In this chapter, my main concern is the question whether we actually treat the fact that we have an intuition (or the intuiting) as evidence in philosophy. For simplicity, I will sometimes use 'an intuition' for 'the fact that we have an intuition'.

Some philosophers are reluctant to use the word ‘intuition’, since they do not think that intuitions are epistemically or ontologically different from other kinds of judgments or beliefs.¹ Although in the last chapter I defended what I called a ‘non-rationalist inflationist’ account of intuitions according to which intuitions are appearance states, nothing in this chapter presupposes my view. For our purposes, it does not matter how we label the respective attitude, and I will mostly stick to the word ‘intuition’—thereby talking about the intuition, judgment, inclination to judge, belief, or whatever else philosophers think intuitions are. The question whether we rely on the fact that we have an intuition as evidence is independent of the question of what intuitions are.

Why should we assume that the fact that we have an intuition (or the intuiting) can or should be used as evidence in philosophy at all? Here are some preliminary reasons. First, Weatherson [132] argues that experts in other areas (e.g., in business) sometimes use the fact that they have an intuition as evidence. Surely we should not take just any intuitions into account, we should rather carefully choose the ones we have reason to think are reliable. But there is no reason to think that philosophy is in any important way different from other disciplines where we rely on experts’ intuitions. It is, however, controversial whether philosophers are experts in a sense that would make them better intuiters than lay people. I present arguments in favour of the claim that philosophers are expert intuiters in section 3.6.

Second, many contemporary philosophers engage in practices which suggest that they take the fact that we have an intuition as evidence. For instance, many philosophers aim to give pragmatic or psychological explanations for the fact that we have a certain intuition that P if this intuition is not treated as true by their theory. I discuss examples of this practice in Chapter 4. Assuming that our practice in philosophy is more or less adequate, this suggests that we can use the fact that we have an intuition as evidence.

¹E.g., Deutsch [28], Earlenbaugh & Molyneux [31]. Williamson [138, p. 152] pleads for the elimination of the term ‘intuition’.

Third, some philosophers think that the fact that we have a sensory seeming is evidence, and ways of defending this view can equally be applied to intuitions. Jessica Brown [10] considers three ways to do so within an internalist epistemology: dogmatism (e.g., Pryor [95]), *a priori* entitlement (e.g., Wright [142]), and explanationism (e.g., Vogel [128]). Brown also discusses an externalist approach she favours, according to which our beliefs about the world are justified if they are reliably formed on the basis of the relevant facts about our psychological states. Assuming that forming the belief that *P* on the basis of the fact that we have an intuition that *P* is reliable, this belief is thereby justified.² I discuss this argument in Chapter 5.

In what follows, I will argue that two ways of undermining the experimental philosophers' challenge to our use of intuitions in philosophy fail.

3.2 Undermining the Experimentalists' Challenge?

I will first give a more detailed outline of the two above-mentioned attempts to undermine the experimentalists' challenge, which I will then address individually in the succeeding sections. Deutsch [28], [29] and Earlenbaugh & Molyneux [31] defend the idea that we do not appeal to intuitions as evidence in philosophy. If this were true, it would be hard to see how experimental philosophy would pose a challenge to philosophical methodology.

Both attempts to undermine the experimentalists' challenge can be understood in the spirit of Timothy Williamson [140], who argues that appealing to psychological states such as intuitions as evidence is an unnecessary practice we should not pursue. Williamson claims that philosophers who appeal to intuitions as evidence do so as a consequence of the misguided idea of *Evidence Neutrality*.³ *Evidence Neutrality* is the thesis that

[...] whether a proposition constitutes evidence is in principle un-

²Brown [10, pp. 506-515].

³Williamson [140, pp. 208-215]. Williamson [138] gives an alternative explanation as to why some philosophers appeal to intuitions, which I discuss in Chapter 5.

contentiously decidable, in the sense that a community of inquirers can always in principle achieve common knowledge as to whether any given proposition constitutes evidence for the inquiry. [140, p. 210]

In cases where philosophers disagree with their peers on whether a proposition about the world is evidence, they draw on a fact about which mutual agreement is easier to achieve, which is the fact that someone has an intuition.⁴ The fact that someone has an intuition then is supposed to count as common evidence. Williamson uses the Gettier Cases as an example:

Arguing from the Gettier proposition that the subject in a Gettier case lacks knowledge, I conclude that knowledge is not equivalent to justified true belief. Now I meet someone who thinks the Gettier proposition [is] a mere cultural prejudice, not itself evidence. In this context, it is not in principle uncontentiously decidable that the Gettier proposition is evidence. Thus the only way to satisfy Evidence Neutrality is by ruling that the Gettier proposition does not constitute evidence. To argue that knowledge is not equivalent to justified true belief, I must go back a step to less contentious premises. What can they be? My opponent allows that I *believe* the Gettier proposition, and may even admit to feeling an inclination to believe it too (I am not merely idiosyncratic), while overriding it on theoretical grounds. [140, p. 211]

However, Williamson thinks that appealing to psychological states as evidence is a practice we should not pursue.⁵ The reason is that there is a gap between psychological states with certain contents and the truth of these contents which is ‘not easily bridged’ and which provokes scepticism. Attempts to psychologise

⁴According to Williamson [138], intuitions are just beliefs or inclinations to believe, on which see Chapter 2.

⁵Williamson does not say that facts about psychological states can never be used as evidence in philosophy. However, he thinks that they should not be used as evidence in the case of thought experiments as counterexamples to philosophical theories, see [140, chapter 7].

the matter of philosophy (e.g., by holding that we are concerned with our concepts only) do not solve the problem, since ultimately, we are interested in facts about the world and not in facts about our psychological states. We should therefore give up *Evidence Neutrality* and appeal directly to facts about the world as evidence. I discuss Williamson's arguments to the effect that we ought not appeal to intuitions as evidence at length in Chapters 4 and 5.

In very much the same spirit, Deutsch argues that we usually or at least in frequently discussed cases do not treat intuitions as evidence. Contrary to the common perception and to what Machery, Mallon, Nichols, & Stich [81] claim, Deutsch thinks that Kripke does not appeal to intuitions as evidence in the Gödel Case, but to facts about the world and to arguments instead.⁶ Here is how Deutsch describes the case:

Kripke offers direct arguments against [...] [the] descriptivist theory of meaning, but he also objects to it indirectly by criticizing the theory of reference it entails. *D* encapsulates the theory of reference that is a consequence of the descriptivist theory of meaning:

D: An ordinary proper name, *n*, as used by a given speaker, *S*, refers to the object that is the denotation of some/most/all of the definite descriptions *S* associates with *n*.

To show that *D* is false, Kripke simply describes counterexamples—cases in which a name, as used by a given speaker, does not refer to the denotation of the definite description(s) the speaker associates with the name. Here is one such case, one of Kripke's own (Kripke, 1972/1980, pp. 83-84): Imagine that Gödel did not prove the incompleteness of arithmetic but that some other man, Schmidt, did. Gödel stole the proof from Schmidt and published it under his own name. But now imagine a speaker who uses 'Gödel', but associates

⁶Deutsch [29] also discusses Weinberg, Nichols, & Stich's [135] criticism of the claim that intuitions from the Gettier Cases play a role as evidence.

just a single description with it, namely ‘the prover of incompleteness.’ To whom does this speaker’s uses of ‘Gödel’ refer, Gödel or Schmidt? The answer, Kripke says, is Gödel, not Schmidt. If Kripke is right, *D* is false. [28, p. 446]

According to Deutsch, ‘nothing in Kripke’s famous argument against the descriptivist theory of reference for proper names hinges on assuming anything about peoples’ intuitions’.⁷ Deutsch concludes that whether a counterexample is intuitive might be psychologically interesting; what matters philosophically is whether it is genuine and hence refutes the respective philosophical theory.

Even more radically, Earlenbaugh & Molyneux argue that there simply is no such practice of treating intuitions as evidence in philosophy and that intuitions only play a rhetorical role in that they function persuasively and give rise to beliefs. Earlenbaugh & Molyneux contrast what they call the ‘evidential-role view’, which is the view that intuitions are used as evidence in philosophy, with what they call the ‘evidential view’, which is the thesis that intuitions in fact are evidence in philosophy.⁸ Whereas experimental philosophers claim that intuitions play an evidential-role but deny the evidential view, Earlenbaugh & Molyneux hold that intuitions do not, from the beginning, play an evidential-role at all. They introduce their project as follows:

We argue that intuitions *do not play an evidential-role in philosophy*. Hence, we show that any evidential view of intuitions that is motivated by the way intuitions are actually used in philosophy is wrongheaded. This sets us aside, we think, from the traditional debate, in which one side argues that they play an evidential-role and so they should, because they are genuine forms of evidence, whereas the other argues that they do, but they should not, because they are not. We argue that they do not, whether or not they should.
[31, p. 92]

⁷Deutsch [28, p. 445].

⁸Earlenbaugh & Molyneux [31, p. 92].

In what follows, I first show that Earlenbaugh & Molyneux's arguments against the evidential-role view of intuitions fail. The authors do not succeed in showing that we do not use intuitions as evidence in philosophy (section 3.3). I argue against Deutsch that he uses the wrong criterion to decide whether Kripke appeals to an intuition in his argument (section 3.4). I then show that we may sometimes rely on intuitions as evidence even if we do not appeal to intuitions (section 3.5).

Having shown that the authors fail to establish that intuitions do not play a role as evidence in philosophy, we have to face the experimentalists' challenge that intuitions are not reliable. In the last part of the chapter I present some arguments to the effect that scepticism about the reliability of intuitions is not warranted (section 3.6).

3.3 Treating Intuitions as Evidence

In order to establish that we do not treat intuitions as evidence, Earlenbaugh & Molyneux give two arguments, one against the claim that intuitings (the propositional attitudes) play an evidential-role and one against the claim that intuiteds (the contents of the propositional attitudes) play such a role. To show that intuitings do not play an evidential-role, Earlenbaugh & Molyneux introduce the following test:

One can check whether a community treats a mental state ψ as a basic evidential state by determining whether the members of that group are willing to accept, *prima facie*, inferences of the following sort:

S ψ 's that P

Therefore P.

[31, p. 98]

‘Basic evidential states’ are states such that a subject *S*’s being in such a state is *prima facie* evidence for the truth of its content. Examples are sensory seemings, memory states, and introspections.⁹ According to Earlenbaugh & Molyneux, it is easy to see that basic evidential states pass the above test:

One is usually willing to infer P from the fact that S seems to see that P, *prima facie*, provided that one believes that S has a normal, functioning visual system, that conditions are normal, and that nothing else is amiss (e.g., one does not believe S to be a liar).

[31, p. 98]

The same is true for any ψ that is a basic evidential state, no matter who has it. In contrast, we are not willing to infer P from the fact that someone different from ourselves has an intuition that P. Intuitings therefore do not pass the test and hence are not treated as basic evidential states, as Earlenbaugh & Molyneux claim:

We clearly hold nothing like GI:

GI: In general, intuitions (intuitings) are reliable indicators of the truth of their contents.

For if we did hold something like GI, then inferences like the following:

Jones has the intuition that P.

Therefore P.

would seem as solid, *prima facie*, as the corresponding inference from one’s own intuitions. But they do not. This shows that we, in practice, are not committed to anything like GI. In fact, we are hardly affected by the intuitions of others, *no matter how strongly*

⁹See also Goldman & Pust [46, p. 180].

they are professed and no matter how many others have the intuition. Thus we are not treating intuitions as basic evidential states.

[31, p. 99]

To reach their conclusion, the authors make two different assumptions. First, they take it to be obvious that whereas we treat basic evidential states as evidence no matter which honest person with a normal, functioning perceptual system has them, we only treat our own intuitions as evidence. Second, they assume that a kind of state is only treated as a basic evidential state if all tokens (where the subject is honest and otherwise normal) of this kind of state are treated as evidence. I think that both assumptions are false.

As to the first assumption, it is not true that we only trust our own intuitions. We certainly trust *some* other people's intuitions. There is no reason to think that students, when they learn philosophy, merely trust philosopher's arguments and do not trust any of the intuitions their teachers communicate. Moreover, I think we would misdescribe what we do in philosophy if we thought that professional philosophers only trusted their own intuitions. We often adopt a philosopher's theory without going through every single step ourselves if we know that this other philosopher is an expert in the respective area and has examined more theories and has more knowledge than we ourselves have. However, it seems true that we trust our own intuitions as evidence if we ourselves are experts in a certain area and someone else, who is an expert in the same area, has a contrary intuition. But the fact that in some cases we prefer to trust our own intuitions as *prima facie* evidence does not imply that we do not sometimes trust other people's intuitions. Intuitions are not subject sensitive in the way Earlenbaugh & Molyneux suggest.

As to the second assumption, Earlenbaugh & Molyneux claim that because we do not trust just anyone's intuitions, we do not in general trust intuitions. However, it seems that we do not trust just anyone's sense of hearing, taste, or vision either. In numerous cases we only trust experts' sensory seemings. For instance, if I want to know which bird is tweeting in my garden, I ask an

ornithologist rather than a friend who has read a beginners book on birds and how they tweet. Only what the expert seems to hear is relevant in this case. Nevertheless, we would not want to say that my friend's sense of hearing is not normal. Think of a geologist who is able to distinguish between two very similar kinds of minerals, or think of a wine taster. We would certainly trust the geologist or the wine taster on certain matters rather than anyone else, but we would not want to say that everyone else's senses are not normal. As in the case of intuitions, experts probably only trust themselves in some cases. If two expert ornithologists do not both seem to hear the same mockingbird tweet (for instance, one of them seems to hear a Chilean Mockingbird, the other a Bahamas Mockingbird), they will probably each trust themselves and not the other expert. It seems that whose sense of hearing, taste, or vision we trust does not depend on the general reliability of their faculty but on other, more specific skills which makes some of us experts. There is no categorical difference between how we treat intuitions and how we treat basic evidential states. If we assume that a kind of state is only treated as a basic evidential state and hence as evidence if all tokens (where the subject is honest and her senses are normal) of this kind of state are treated as evidence, there are probably no states that we treat as basic evidential states.

Even though in this chapter I am mainly concerned with intuitings as evidence, I will briefly look at Earlenbaugh & Molyneux' second argument, the argument against the evidential-role of *intuiteds*. The authors claim that intuiteds do not play what they call a 'psycho-evidential role', and this is why we do not treat them as evidence. Propositions play a psycho-evidential role if they play a role as evidence in virtue of being the target of a psychological state.¹⁰ According to Earlenbaugh & Molyneux, intuiteds in general do not play a psycho-evidential-role because we do not *necessarily* place credence in our intuiteds:

Believed propositions play a psycho-evidential-role because the be-

¹⁰Earlenbaugh & Molyneux [31, p. 100].

lief guarantees that the subject invests credence in them. But S's intuiting P is not sufficient for S investing credence in it. Thus, if an inference legitimately proceeds from an intuited proposition, it must be because the subject stands in some other, credence-entailing, attitude towards it. [31, p. 102]

Obviously, we do not place any credence in the contents of intuitions we do not trust (e.g., folk intuitions). Moreover, even if we usually trust our own intuitions, we do not necessarily place any credence in their content. In cases where we firmly believe that *not-P*, we do not place credence in the content of our intuition that *P*. Take Jackson's Mary Case. Suppose that *S* holds that physicalism is true. *S* then thinks about the Mary Case for a long time, and eventually comes to the conclusion that she was right and physicalism is indeed true. Now suppose furthermore that whenever *S* goes through the Mary case, she nevertheless has the intuition that Mary learns something new when she leaves her black-and-white room. *S* firmly believes that physicalism is true, and she knows that her intuition about Mary is incompatible with the claim that physicalism is true. In this case, it is obvious that *S* does not place any credence in the content of her intuition. The evidence that *S* gets from the fact that she has an intuition is defeated by the overwhelming evidence she has for the truth of physicalism.

As Bealer (e.g., [7]) and others have discussed in their work, this phenomenon is well-known from cases of optical illusions.¹¹ Since we know that a straight stick in water is not bent, we do not place any credence in the content of our perceptual seeming that it is bent. Similarly, because we know that the two lines in the Müller-Lyer illusion are of equal length, we do not place any credence in the content of our perceptual seeming that they are not of equal length. However, that in some cases we do not place any credence in their contents does not mean that we think sensory seemings do not play a psycho-evidential role. The same should obviously hold for intuitions.

¹¹See also Pust [96].

Let me make a last general remark on Earlenbaugh & Molyneux’s arguments against the idea that we treat intuitions (intuitings or intuiteds) as evidence. The authors admit that it *seems* as if we treat intuitions as evidence, which they explain by claiming that intuitions cause and motivate beliefs. According to their use of the phrase ‘treat as evidence’, a subject *S* can appeal to an intuiting or intuited as evidence and believe that it is evidence without thereby treating it as evidence. Whether we treat a particular intuiting as evidence depends on whether we treat *all* (relevant) intuitings as evidence, and whether we treat a particular intuited as evidence depends on whether we place credence in *all* (relevant) intuiteds. This is a counterintuitive use of the word ‘treat’. We would like to say that *S* treats an intuition as evidence when *S* intends to use the intuiting or intuited as evidence and behaves as if it was evidence. In what follows, this is how the word ‘treat’ will be used in this chapter.

3.4 Referring to Intuitions

In order to show that Kripke appeals to facts about the world as evidence in the Gödel Case, Deutsch points out that Kripke does not say it is *intuitive* that we are not talking about Schmidt. Instead, Kripke says ‘straight out, and emphatically, that we are not talking about Schmidt’.¹² According to Deutsch, this indicates that Kripke does not treat intuitions as evidence. Here is the relevant passage in Deutsch which he uses to defend the claim that Kripke’s argument does not appeal to intuitions:

In his own discussion of the case, Kripke, after spinning the tale of Gödel and Schmidt, and using ‘we’ to refer to those of us who, in the story, associate just ‘the man who discovered the incompleteness of arithmetic’ with ‘Gödel,’ says that, on descriptivism, since the man who discovered the incompleteness of arithmetic is in fact Schmidt, we, when we talk about ‘Gödel’, are in fact always refer-

¹²Deutsch [28, p. 451].

ring to Schmidt' (Kripke, 1980, p. 83). Immediately following this comment, Kripke says, 'But it seems to me that we are not. We simply are not' (Kripke, 1980, p. 84). He does not say that it is *intuitive* that we are not talking about Schmidt; he says straight out, and emphatically, that we are not talking about Schmidt. [28, p. 451]

Deutsch gives two reasons for thinking that his view is correct. First, he shows that Kripke *argues* for his view and does not simply rely on his intuition that 'Gödel' does not refer to Schmidt. Second, Deutsch claims that there is a better reconstruction of Kripke's argument without reference to an intuition, which corresponds with what Kripke literally says. Here is the form of the argument Deutsch thinks is not correct:

- (0) It is intuitive that there is an F that is not a G .
- (1) So, there is an F that is not a G .
- (2) Hence, not all F s are G s.

Instead, the best reconstruction of the argument is this:

- (1) There is an F that is not a G .
- (2) Hence, not all F s are G s.¹³

I think that both ways of arguing do not show what Deutsch wants to establish. First, that Kripke gives additional arguments for his view does not imply that he does not use his intuition as evidence. Second, the best reconstruction of an argument is not necessarily the correct reconstruction of an author's reasoning. Surely, once we have all the evidence for the claim that 'Gödel' does not refer to Smith from Kripke's arguments, we do not need to use the initial evidence in our reconstruction anymore. However, this does not mean that Kripke did not appeal to an intuition as initial evidence when he first came up with his counterexample to the descriptive theory of meaning.

¹³Deutsch [29, p. 452].

While I think it is very plausible that Kripke appeals to an intuition in his argument, my aim is not to establish this interpretation. I will merely argue that Deutsch does not succeed in showing the contrary, because the evidence he gives does not favour the claim that Kripke does not appeal to an intuition over the claim that he does.

To see this, consider the following. If we want to find out whether we appeal to intuitions in philosophy, at first sight, the obvious question to ask seems to be the following.

(a.) Do we use the term ‘intuition’ in the premises of our arguments?

(a.) can easily be answered by counting the word ‘intuition’ in philosophy papers. The result will be that in contemporary analytic philosophy, the word is used quite frequently, but possibly not so in the history of philosophy. However, it does not follow from the fact that we use the word ‘intuition’ more frequently in contemporary philosophy that we necessarily appeal to intuitions as evidence more frequently. Maybe we sometimes use the word as a hedging term or as an expression of modesty, etc. Maybe there is no psychological state that corresponds to our use of the term, which could mean that we never actually use ‘intuition’ to refer to evidence.¹⁴

Deutsch points to the fact that Kripke does not use ‘intuition’ as evidence for the claim that Kripke’s argument does not involve intuitions. However, as we cannot infer simply from the fact that we use the word ‘intuition’ that we actually refer to a psychological state, we cannot infer from the fact that we do *not* use the word ‘intuition’ that we do *not* refer to an intuition. It could be that we refer to a psychological state with some expressions other than the term ‘intuition’. The point is that (a.) is not the adequate criterion to decide whether someone appeals to an intuition as evidence or not. Whether

¹⁴See, e.g., Cappelen [12].

we use the term ‘intuition’ does not show anything. We have to ask two further questions:

(b.) Do we, in the premises of our arguments, use the term ‘intuition’ to refer to a psychological state that serves as evidence for our beliefs?

(c.) Do we, in the premises of our arguments, refer with some other expressions to a psychological state that serves as evidence for our beliefs?

(b.) and (c.) are probably the most discussed questions concerning intuitions in philosophy. As to (b.), Deutsch thinks that when we say that p is intuitive, this means something like ‘we know p directly, without inference’, ‘ p is true’, ‘ p is pretty obviously true’, or ‘ p should be accepted as true unless compelling reasons can be given for rejecting it’.¹⁵ Williamson [138] [140], and others argue that what we call ‘intuitions’ can be reduced to beliefs or belief-like states. They think that the content of the intuition can be evidence, but that there is no special psychological state such that the fact that we have it could or should be used as evidence. Others, like Bealer [7], Pust [96], Sosa [115], and Weatherson [132] argue that the fact that we have an intuition can be evidence in philosophy. I agree with the latter that we refer to a psychological state when we use the word ‘intuition’ in philosophy (at least in paradigm cases such as thought experiments), and I also think that we can appeal to the fact that we have an intuition as evidence, as I mentioned in section 3.1 (and as I defend in Chapters 4 and 5). However, each of these answers to (b.) need to be defended, and merely looking at our use of the word ‘intuition’ does not do the work.

Complementarily, we have to ask whether there are cases where we appeal to intuitions as evidence without using the word ‘intuition’ (c.). Some philosophers (traditional philosophers as well as experimental philosophers)

¹⁵Deutsch [28], [29, p. 458].

have understood the Gödel Case and similar cases such that intuitions play an evidential role, even though their authors do not use the term ‘intuition’. ‘It seems to me that we are not’ in Kripke’s argument might, after all, refer to an intuition. This terminology has been used by defendants of rationalist accounts of intuitions such as Bealer [7], who takes intuitions to be *intellectual* seemings in contrast to sensory seemings.¹⁶ Whether Kripke refers to an intuition with the phrase ‘It seems to me’ cannot be decided simply by looking at what Kripke literally says. The words we use might be *some* evidence as to what we refer to, but it is clearly not enough evidence in a philosophical debate like the one over intuitions. We can decide on this matter only on a theoretical level.

3.5 Intuitions as Implicit Evidence

Let us, for a moment, go back to Earlenbaugh & Molyneux [31]. They argue that intuitings do not play any role as evidence in philosophy because we do not treat them as basic evidential states. We do not treat intuitings as basic evidential states because we do not treat all (relevant) intuiting tokens as evidence, which the authors think is necessary for a kind of state to be treated as a basic evidential state. They draw the following conclusion concerning the role of intuitions as evidence:

[...] if intuitions were merely evidence but [...] were not treated as such then they would play little role in argumentation and theory construction. They would be inert – indicative of the truth but not in a way to which we have access. Hence, intuitions can explain the behaviors of the philosophical community only if they are treated as evidence; simply being evidence is not enough. [31, p. 92]

Earlenbaugh & Molyneux claim that as far as they know, there is no motivation for a view according to which intuitions are not treated as evidence but

¹⁶See Chapter 2.

still play a role as evidence.¹⁷ Given what they mean by ‘treated as evidence’ (see section 3.3), this might be plausible. However, there is motivation for such a view if we take ‘treating intuitions as evidence’ to mean ‘appealing to intuitions as evidence’. I will now argue that in order for the fact that we have an intuition to be our evidence, we do not necessarily have to appeal to it.

Brian Weatherson [132] has argued that a view according to which the fact that we have an intuition that P is evidence in cases where P is false can give us an explanatory advantage. Take, for instance, Frank Jackson on his Mary Case against physicalism. Suppose, for the sake of the argument, that with his sentence ‘It seems just obvious that she will learn something about the world and our visual experience of it’¹⁸, Jackson did not refer to the fact that he had an intuition but rather to the putative fact that a person in Mary’s situation learns something new. Suppose that he also did not believe that the fact that he had an intuition was his evidence. Maybe he thought that only the contents of our intuitions can possibly be evidence in philosophy. Suppose furthermore that Jackson *wrongly* thought that a person in Mary’s situation learns something new. In fact, Mary does not learn anything new and physicalism is true.¹⁹

According to a Williamsonian view of evidence, the putative fact that a person in Mary’s situation learns something new could not have been Jackson’s evidence. In his book *Knowledge and its Limits*, Williamson [137] famously defends his view that all and only knowledge is evidence. As a consequence, evidence is factive and propositional, and even though we might treat false propositions as evidence sometimes, only true propositions actually are evidence.²⁰ According to such a view, we might want to say that Jackson did not have any evidence from the thought experiment which counts against physicalism. In this case, we might want to call his belief that physicalism is false

¹⁷Earlenbaugh & Molyneux [31, pp. 92-93, footnote 8].

¹⁸Jackson [54, p. 182].

¹⁹Most philosophers think that the Mary Case is not a genuine counterexample to physicalism, including Jackson [55].

²⁰See Williamson [137, p. 201].

in some sense irrational (supposing that the thought experiment was his only source of evidence against physicalism).²¹

However, we might as well say that Jackson relied on some genuine evidence for his belief, namely the fact that he had the intuition that Mary learns something new, even though he did not appeal to this fact as evidence. According to Williamson [137], we do not always know what our evidence is, and there are two ways in which a belief can be based on evidence:

Call one's belief in *p* *explicitly* evidence-based if it is influenced by prior beliefs about the evidence for *p* [...] Call one's belief in *p* *implicitly* evidence-based if it is appropriately causally sensitive to the evidence for *p* [...] the causal sensitivity of the belief in *p* to the evidence for *p* need not be mediated by further beliefs about the evidence for *p*. [137, pp. 191-192]

Jackson's belief that physicalism is false could have been appropriately causally sensitive to the *implicit evidence* consisting in the fact that Jackson had an intuition that Mary learns something new.

This explanation as to how Jackson's belief was based on evidence does not work if we endorse what Goldman & Pust [46] have suggested, which is a 'two-step evidential route' for evidence from thought experiments:

An adequate reconstruction of philosophical methodology [...] requires a two-step evidential route. In the first step, the occurrence of an intuition that *p*, either an intuition of one's own or that of an informant, is taken as (prima facie) evidence for the truth of *p* (or the truth of a closely related proposition). In the second step, the truth of *p* is used as positive or negative evidence for the truth of a general theory. [46, p. 182]

²¹Weatherson [132] argues that philosophers sometimes have false intuitions, but we would not want to call them irrational for this reason. If we want to call them rational and if being rational requires respecting the evidence, we have to say that they have genuine evidence.

According to Goldman & Pust, the fact that we have an intuition that P is evidence for the fact that P , which is then evidence against Q . Applied to the Gettier Cases, the fact that we have the intuition that a subject in a Gettier situation does not have knowledge is first taken as *prima facie* evidence for the claim that a subject in such a situation would not have knowledge. In a second step, that such a subject lacks knowledge is then taken as evidence against the JTB theory of knowledge, i.e., for the claim that the JTB theory is false. Applied to the Mary Case, the fact that Jackson had the intuition that Mary learns something new is *prima facie* evidence for the claim that a person in Mary's situation learns something new. Since, however, a person in Mary's situation does not learn anything new, we have no evidence against physicalism.

However, a different understanding of the evidential route according to which there is only one step seems to be more widely held. When S has the intuition that P , S has evidence for P and at the same time evidence against Q . This is plausible since P is supposed to imply *not-Q*. Applied to the Mary Case, the fact that Jackson had an intuition that Mary learns something new was evidence not only for the claim that a person in such a situation would learn something new, it was also evidence against physicalism, since the first implies the second. Such a view is especially held by philosophers who think that the only evidence we gain from thought experiments against the target theory is the fact that we have an intuition. I will discuss the question of what exactly the evidence is we gain from thought experiments in Chapter 4.

I have argued that not only can we refer to an intuition without using the word 'intuition', we can also rely on an intuition as evidence without explicitly appealing to it. The role intuitions play as evidence can neither be simply read off our words nor off the premises of our arguments. I therefore take it that the experimental philosophers' challenge cannot easily be undermined and has not been undermined by Deutsch's or Earlenbaugh & Molyneux's arguments.

3.6 The Expertise Defence

If the experimental philosophers' challenge cannot be undermined, we have to face it. In what follows, I will present reasons to think that empirical findings about the unreliability of intuitions are not relevant to philosophers' intuitions and scepticism about the reliability of intuitions in philosophy is not warranted. This line of argument has been brought forward, amongst others, by Williamson [141]. Even though Williamson himself argues that we ought not appeal to the fact that we have an intuition, his argument for the irrelevance of experimental philosophers' findings applies to both the content of the intuition and the fact that we have an intuition.

Experimental philosophers have tested laypeople's intuitions on philosophical questions.²² Their concern is that philosophers might be prone to the same kinds of effects and psychological features they have discovered laypeople to be. Some experimental philosophers draw a radical conclusion from their studies: using intuitions as evidence for or against philosophical theories is an unreliable method which should not be pursued (e.g., Alexander & Weinberg [2], Weinberg [133]).

However, one might wonder what exactly the conclusions are we should draw for philosophy from the results of testing laypeople's intuitions. Some philosophers, e.g., Devitt [30] and Goldman & Pust [46], are interested in laypeople's intuitions because they think that only laypeople's intuitions are uncontaminated by philosophical theory. Goldman & Pust say the following:

If the person experiencing the intuition is a philosophical analyst who holds an explicit theory about the nature of F, this theory might warp her intuitions about specific cases [...] For this reason, philosophers rightly prefer informants who can provide pre-theoretical intuitions about the targets of philosophical analysis, rather than informants who have a theoretical 'stake' or 'axe to

²²These laypeople are mostly undergraduate students in the United States.

grind'. [46, p. 83]

Other philosophers such as Hilary Kornblith [65], Kirk Ludwig [78], and Williamson [140] think that experimental philosophers' studies are irrelevant because we are interested in expert intuitions only rather than in laypeople's intuitions. These philosophers take it that philosophers' intuitions in thought experiments are more reliable than laypeople's intuitions because like experts in other disciplines, philosophers have specific skills and specific philosophical knowledge and experience which makes them better intuiters.²³ Kornblith claims that like in the sciences, where 'accurate theory is allowed to play a role in guiding and shaping observation', we should not prefer 'the intuitions of the ignorant and the naive over those of responsible and well-informed investigators'²⁴. Williamson [140] expresses it thus:

We should not regard philosophical training as an illegitimate contamination of the data, any more than training natural scientists how to perform experiments properly is a contamination of their data. Although the philosophically innocent may be free of various forms of theoretical bias, just as the scientifically innocent are, that is not enough to confer special authority on innocent judgment, given its characteristic sloppiness. [140, p. 191]

Weinberg, Gonnerman, Buckner & Alexander [134] have criticized the so-called expertise defence and argued that whether philosophers are experts in the relevant sense is an empirical question. They claim that the burden of proof is on the traditional philosophers to show that their intuitions are in fact more reliable. They moreover argue that there is little evidence that philosophical training is of any relevance to a reliable performance in thought experiments. Philosophers should therefore not rely on intuitions from thought experiments before having conducted the relevant empirical investigations.

²³Ludwig [78] compares philosophers to mathematicians, Williamson [139], [140] to lawyers, and Weatherston [132] to business people.

²⁴Kornblith [65, pp. 33-34].

In reply, Williamson [141] argues that the burden of proof is rather on the experimental philosophers' side to show empirically that philosophical training is not conducive to a reliable performance in thought experiments. He argues that philosophers' training fulfills all relevant criteria mentioned in the literature on expertise: (a) repetitive practice with fast, accurate feedback, (b) decomposition of the task into sub-tasks, and (c) use of external decision aids. Philosophers often get immediate feedback on thought experiments (a), running a thought experiment involves several steps and different skills (b), and formal methods can help us to evaluate our judgments in thought experiments (c). Williamson concludes that Weinberg, Gonnerman, Buckner & Alexander

[...] provide no significant evidence that thought experimentation is worse off in the relevant respects than the cognitive skills they acknowledge to be enhanced by training in philosophy, such as informal argumentation and the close analysis of texts. Consequently, they provide no reason to rely less on trained philosophers' skill at thought experimentation than on their skill at those other cognitive tasks. [141, p. 226]

For the sake of my arguments throughout this thesis, I will assume that Williamson and others are right: in general, philosophers' intuitions on philosophical matters are reliable.

This, however, does not imply that philosophers' intuitions are infallible, and it leaves room for empirical or other explanations as to why we have a particular intuition. Philosophers might sometimes be prone to unwanted psychological effects, cultural biases and pragmatic phenomena. For this reason, it is important to take empirical research into account in order to decide whether we can trust a particular intuition on a particular subject matter.

In some experimental philosophers' work, a more constructive claim has been made, namely that empirical findings can supplement and improve our methods in philosophy (e.g. Knobe [61], Knobe & Nichols [62], Swain, Alexan-

der, & Weinberg [119], Nadelhoffer & Nahimas [85]). James Beebe gives the following characterization of empirically informed philosophy:

Experimental epistemology is the use of the experimental methods of the cognitive sciences to shed light on debates within epistemology, the philosophical study of knowledge and rationally justified belief. Some skeptics contend that ‘experimental epistemology’ (or ‘experimental philosophy’ more generally) is an oxymoron. If you are doing experiments, they say, you are not doing philosophy. You are doing psychology or some other scientific activity. It is true that the part of experimental philosophy that is devoted to carrying out experiments and performing statistical analyses on the data obtained is primarily a scientific rather than a philosophical activity. However, because the experiments are designed to shed light on debates within philosophy, the experiments themselves grow out of mainstream philosophical debate and their results are injected back into the debate, with an eye to moving the debate forward. This part of experimental philosophy is indeed philosophy—not philosophy as usual perhaps, but philosophy nonetheless. [8, p. 1]

Such a positive view of experimental philosophy and the relevance of empirical experiments to philosophical questions has been practiced, amongst others, by Jennifer Nagel [86], [87] and Tamar Gendler [40], [39].

In the above quote, Beebe suggests that some parts of experimental philosophy is philosophy, while engaging in experiments is science. But we might as well take the empirical work to be part of philosophy, and we can then define philosophy via the central questions philosophers have been trying to answer rather than via the methods philosophers use to answer these questions. This is the view I will hold in what follows. While I take it that scepticism about the reliability of intuitions in philosophy is not warranted, I think that psychology

and experimental philosophy can sometimes help us to answer philosophical questions.

3.7 Conclusion

In this chapter, I argued that two attempts to undermine the challenge experimental philosophy poses to the role of intuitions in philosophy are not successful. I first argued that Earlenbaugh & Molyneux's arguments against the evidential-role of intuitions fail and that Deutsch does not use the correct criterion to decide whether Kripke or anyone else refers to intuitions in their arguments. I then showed that even if we do not appeal to the fact that we have an intuition as evidence from thought experiments, this fact can constitute implicit evidence against the target theory. I finally presented reasons to believe that scepticism about the reliability of intuitions as advocated by experimental philosophers is not warranted, and I mentioned a more positive picture according to which experimental philosophy and psychology can help us to answer philosophical questions.

Whereas Kripke and others do not use the word 'intuition', it seems that using it and also appealing to intuitions as evidence has become fashionable in recent years. In the following Chapter 4, I argue that the practice of appealing to the fact that we have an intuition is wide-spread, and I criticize Williamson's view that it is a practice we ought not pursue. In Chapter 5, I offer an explanation as to why we appeal to intuitions as evidence from thought experiments according to which this is a good practice.

Chapter 4

Thought Experiments and Evidence

Abstract

In this chapter, I am concerned with the kind of evidence from thought experiments we ought to appeal to. I defend the view that we ought to appeal only to the fact that we have an intuition, and that this evidence is *initial* evidence only. I first show that this corresponds with our practice. I then argue that Timothy Williamson's [138] worry that appealing to the fact that we have an intuition as evidence from a thought experiment leads to a regress of giving evidence for evidence does not apply to this view. Appealing to the fact that we have an intuition as initial evidence allows us to stop the regress and is just one way of pursuing a practice that is commonly performed in the sciences as well.

4.1 Introduction

Timothy Williamson [140, chapter 6] gives the following account of thought experiments. Take any Gettier Case as a counterexample to the theory of knowledge as justified true belief (the JTB theory) as an example. Let x be a person in the Gettier Case and p the relevant justified true proposition. GC is the Gettier relation which holds between x and p .

1. $\diamond \exists x \exists p GC(x, p)$
2. $\exists x \exists p GC(x, p) \Box \rightarrow \forall x \forall p [GC(x, p) \supset JTB(x, p) \wedge \neg K(x, p)]$
3. $\diamond \exists x \exists p (JTB(x, p) \wedge \neg K(x, p))$

In words, this means that (1) it is possible for some x to stand to some p in a Gettier relation, (2) if some x were to stand to some p in a Gettier relation, she would have a justified true belief which is not knowledge, and (3) it is possible to have a justified true belief which is not knowledge. The antecedent of (2) gives us the judgment we make when confronted with the thought experiment scenario.

(3) creates an inconsistency with the target theory, the JTB theory of knowledge, according to which justified true belief of a proposition p is necessary and sufficient for knowledge of p . By creating this inconsistency, the Gettier Case undermines the JTB theory of knowledge. Hence, the Gettier Case functions as a counterexample to the JTB theory of knowledge.

I will sometimes call the content of the judgment we make when confronted with a scenario in a thought experiment (e.g., that the person in the scenario described does not have knowledge) the ‘thought experiment proposition’. It does not matter how we label the respective attitude towards the thought experiment proposition, but I will mostly stick to the word ‘intuition’—thereby talking about the intuition, judgment, inclination to judge, belief, or whatever else philosophers think the relevant judgments are.

For Williamson, the method of thought experiments is an instance of ordinary counterfactual reasoning. Since our counterfactual judgments are reliable, thought experiments reliably lead to knowledge of the thought experiment proposition. Many philosophers share the view that thought experiments can provide knowledge of the thought experiment proposition, but their explanations as to how this knowledge is gained vary. George Bealer [7], [6] and Ernest Sosa [115] think that intuitions can provide *a priori* knowledge, and since we have an intuition towards the thought experiment proposition,

we can gain *a priori* knowledge of the thought experiment proposition. Like Williamson, Ichikawa & Jarvis [53] think that intuitions are just ordinary beliefs, but they provide an account of thought experiments which entails that thought experiments usually generate *a priori* knowledge. Despite the different views these philosophers take on thought experiments and intuitions, they agree that thought experiments provide a way of gaining knowledge of the thought experiment proposition.

However, since philosophers' intuitions are fallible, it is useful to distinguish between *intuitive* counterexamples and *genuine* counterexamples to philosophical theories. A thought experiment is an intuitive counterexample if most philosophers have an intuition that the thought experiment proposition holds. An intuitive counterexample which undermines the target theory is genuine, and an intuitive counterexample which does not undermine the target theory is merely intuitive.¹ A counterexample is not necessarily intuitive (for instance, there could be empirical counterexamples to philosophical theories), but we are concerned with intuitive counterexamples only.

Note that the Gettier Cases are very convincing counterexamples to the JTB theory of knowledge. Even though philosophers disagree about the correct theory to cover our intuitions in the Gettier Cases, very few doubt that they are genuine counterexamples to the JTB theory and hence that they undermine the JTB theory.² From looking just at the Gettier Cases, it seems that philosophical theories get refuted by the appeal to our knowledge of the thought experiment proposition. When we talk about already established counterexamples like the Gettier Cases, we use the fact that the person in the Gettier Case does not have knowledge in our reconstruction. However, the common use of the Gettier Cases as examples in methodological debates seem to lead to

¹This is in contrast to Weatherson's definition of a counterexample: 'Let us say that a counterexample to the theory that all Fs are Gs is a possible situation such that most people have an intuition that some particular thing in the story is an F but not a G', Weatherson [131, p. 3]. According to my view, this defines an *intuitive* counterexample, but not a *genuine* counterexample.

²Weatherson [131], however, thinks that the JTB theory is still a candidate worth considering.

a distorted view of how theories get refuted in philosophy. Most counterexamples to philosophical theories do not undermine the target theory as obviously as the Gettier Cases undermine the JTB theory.

The issue I address in this chapter concerns our practice of refuting philosophical theories. I defend the view that we ought to appeal only to the fact that we have an intuition, and that this evidence is *initial* evidence only. I first show that this corresponds with our practice. I then argue that Williamson's [138] worry that appealing to the fact that we have an intuition as evidence from a thought experiment leads to a regress of giving evidence for evidence does not apply to this view. Appealing to the fact that we have an intuition as initial evidence allows us to stop the regress and is just one way of pursuing a practice which is commonly performed in the sciences as well.

4.2 Distinguishing Three Questions

In the current debate over the methodology of thought experiments as counterexamples to philosophical theories, the following three questions have not always been distinguished carefully.

- (1) What kind of facts can undermine a philosophical theory?
- (2) What kind of evidence do we gain from a thought experiment?
- (3) What kind of evidence from a thought experiment ought we to appeal to?³

Two different kinds of answers are possible to all three questions: facts about the subject matter the theory is about, or facts about our psychological states with contents about the subject matter the theory is about.

(1) is a question about the ontology of the facts that undermine a philosophical theory. A wide range of contemporary philosophers embrace the view that most philosophical theories are not about our psychological states. Even though philosophical theories usually concern more abstract matters of fact,

³Provided that we gain any evidence against the target theory.

they are as much about the world as scientific theories are.⁴ Accordingly, philosophical theories can be undermined by the facts about the world rather than by facts about our psychological states, i.e., by the facts that are the contents of our intuitions rather than by facts about our intuitions. For instance, our theories of knowledge are about knowledge and not about what we believe knowledge is or what our concept of knowledge is. Accordingly, the JTB theory of knowledge is undermined by the fact that a person in a Gettier Case has justified true belief that p but no knowledge that p rather than by the fact that we have the intuition that a person in a Gettier Case has justified true belief that p but no knowledge that p .⁵ Hence, what ultimately decides whether a theory is correct are the facts about the world, not facts about our psychological states. In what follows, I take (1NP) to be the standard view:

(1NP) Philosophical theories can be undermined by the contents of our intuitions.

There is more disagreement with respect to the answers to questions (2) and (3). (2) concerns the evidence we gain from a thought experiment and hence the epistemic situation we are in with respect to the thought experiment proposition. (3) concerns the methodological advice with respect to the evidence we gain from a thought experiment and with respect to the target theory. By ‘methodological advice’ I mean an answer to the question of what our practice of refuting a philosophical theory ought to be, where ‘to refute’ means to show that the evidence undermines the theory. It seems that especially question (3) has not always been clearly distinguished from (1) and (2).⁶

⁴In the recent debate about thought experiments and intuitions, Goldman & Pust [46] and Goldman [45] are an exception. They think that philosophy is about our concepts rather than about the world.

⁵See, e.g., Kornblith [64, p. 12].

⁶It is not entirely clear whether Williamson distinguishes between an epistemic question (2) and a methodological question (3). Williamson’s 2007 book *The Philosophy of Philosophy* is not supposed to give concrete methodological advice, as the author explicitly claims: ‘I considered using the phrase “philosophical method” in the title, but decided against on the grounds that it seemed to promise something more like a recipe for doing philosophy than I believe possible. When asked for advice on some occasion, the Duke of Wellington is said to have replied “Sir, you are in a devilish awkward predicament, and must get out of it as best you can.” My advice would be scarcely more useful. At the crucial point, I

In the following sections, I first present three possible views resulting from different answers to (2) and (3) (section 4.3). According to the view I defend, the answer to (2) is that both the contents of our intuitions and facts about our intuitions constitute the evidence we gain from a thought experiment, and the answer to (3) is that we ought to appeal to the fact that we have an intuition as evidence against the target theory (section 4.4). To show that this corresponds with our practice, I provide examples of how we discuss intuitive counterexamples to philosophical theories (section 4.5). In the remainder of the chapter, I argue that two worries Williamson raises in his paper ‘Philosophical “Intuitions” and Scepticism about Judgement’ [138] and in *The Philosophy of Philosophy* [140] do not threaten this view. I argue that appealing to the fact that we have an intuition as evidence does not lead to a regress of giving evidence for evidence and hence does not provoke scepticism (section 4.6) and that appealing to the fact that we have an intuition as evidence does not presuppose or lead to a view according to which philosophy is different in principle from the sciences (section 4.7).

4.3 Evidence from Thought Experiments

In this section, I give two views on what our evidence from thought experiments is, and I present Williamson’s reasons to prefer one of them over the other. Williamson draws a picture of the debate according to which many con-

can only say “Use your judgment.” [140, p. ix] Williamson goes on to claim that as the primary task of the philosophy of science is to understand science rather than to give scientists advice, it is the primary task of the philosophy of philosophy to understand philosophy, rather than to give philosophers advice. However, clearly every sub-discipline of science is also concerned with the methodology that gives applicable advice on how to do research. If philosophy is no exception, philosophy needs a sub-discipline that gives methodological advice. Williamson seems to agree at least with the idea that philosophy has methods which can be studied: ‘Some may wonder whether philosophy has a method to be studied, especially if it is as methodologically undistinctive as just suggested. Forget the idea of a single method, employed in all and only philosophical thinking. Still, philosophers use methods of various kinds: they philosophize in various ways. A philosophical community’s methodology is its repertoire of such methods. The word “method” here carries no implication of a mechanically applicable algorithm, guaranteed to yield a result within a finite time. On this loose understanding of what a methodology is, it is disingenuous for a philosopher to claim to have none.’ [140, p. 3]. Moreover, Williamson himself gives advice on certain concrete methodological questions. As I will discuss below, he thinks that we should appeal to the content of the thought experiment proposition rather than to the fact that we have an intuition as evidence from thought experiments, [140, chapter 6].

temporary analytic philosophers think that in thought experiments, ‘[...] our ultimate evidence consists of the psychological truths that we have intuitions [...], whether true or false.’⁷ Whether this claim about most contemporary philosophers is correct or not, the view entails the following answer to (2):

(2P) The evidence we gain consists in the fact that we have an intuition.

From (2P) it follows that the only possible answer to (3) is this:

(3P) We ought to appeal to the fact that we have an intuition.

I will call the view combining (2P) and (3P) the *Psychologistic View* of evidence from thought experiments.

As mentioned in section 4.1, most non-experimental philosophers think that we usually gain knowledge of the thought experiment proposition, because thought experiments reliably generate knowledge. It is also very plausible to suppose that most of the time, we know that we have an intuition (or a belief, etc.) towards the thought experiment proposition. Since according to Williamson [137], all our knowledge constitutes our total evidence, the fact that we have an intuition is most of the time part of our evidence as well. For cases where we know both the content of the intuition and the fact that we have an intuition, this view of evidence leads to the following answer to (2).

(2I) The evidence we gain consists in the content of our intuition and in the fact that we have an intuition.

Williamson [138], [140] thinks that as a general strategy, we ought to appeal to the content of the intuition as evidence, and we should only withdraw this evidence if it turns out that we were wrong and the content of the intuition is false. Williamson claims that this is not only the best strategy in philosophy, it also corresponds with what we do in the sciences:

[...] if we do know that P, would it not be negligent not to use that knowledge in evaluating an empirical theory to which it is relevant?

⁷Williamson [138, p. 119].

It would not advance science to insist that scientists' evidence cannot include the fact that 19 out of 20 rats fed the substance died within 24 hours, but only the fact that the scientist was perceptually inclined to judge that 19 out of 20 rats fed the substance died [...]. Of course, it may later turn out that a disgruntled lab technician fed the rats the wrong substance, but the proper response to such remote possibilities is to backtrack if one of them is found to obtain, not to make a futile attempt in advance to identify evidence for which backtracking will not be required in even the remotest eventualities. [138, p. 151]

Hence, the evidence we ought to appeal to against the target theory consists in the content of our intuition rather than the fact that we have an intuition:

(3NP) We ought to appeal to the content of our intuition.

I will call the view which combines (2I) and (3NP) the *Non-Psychologistic View* of evidence from thought experiments, since the evidence consisting of psychological facts ought to play no role in our practice of refuting a philosophical theory.

Let us have a look at Williamson's reasons against the *Psychologistic View*. The main problem with appealing to the fact that we have an intuition, according to Williamson, is that it is not clear how we get from a truth about our psychology to a truth about the world (which we are ultimately interested in, as stated in (1NP)), since 'psychological evidence has no obvious bearing on many philosophical issues'.⁸

There are three further theses Williamson thinks are closely related to the *Psychologistic View* and more or less explicitly endorsed by a large range of contemporary philosophers. The first thesis is *Evidence Neutrality*, which is the idea that

⁸Williamson [140, p. 234]. See Brown [10] for arguments to the effect that the gap between facts about our intuitions and facts about the world can be closed, see also Chapters 3 and 5. I will discuss this worry in more detail in Chapter 5.

[...] whether a proposition constitutes evidence is in principle uncontentiously decidable, in the sense that a community of inquirers can always in principle achieve common knowledge as to whether any given proposition constitutes evidence for the inquiry. [140, p. 210]

Williamson [140] argues that their false belief in *Evidence Neutrality* leads many philosophers to the view that only the fact that we have an intuition can be evidence from the thought experiment, i.e., to the *Psychologistic View*. The *Psychologistic View* then motivates a second thesis, *Judgment Scepticism*:

Judgment Scepticism: we cannot trust our practices of applying concepts in judgment.⁹

Endorsing *Judgment Scepticism* threatens to lead to an even more general and destructive kind of scepticism which questions our judgments in the sciences as well, such as observational judgments or judgments that rely on folk concepts.

The third thesis Williamson finds to be related to the *Psychologistic View* is that there is a fundamental difference between the sciences and philosophy insofar as we ought to trust our judgments in the sciences, but we ought not trust our judgments in thought experiments.¹⁰ In the context of the debate over the methodology of philosophy, philosophical exceptionalism is mainly associated with the view that intuitions provide or are supposed to provide a special source of knowledge, which equips philosophers with a special kind of method that is fundamentally different from the methods we apply in ordinary thinking or in the sciences.¹¹ For the current purpose, the following claim

⁹Williamson [140, p. 220].

¹⁰Williamson [140, p. 220].

¹¹Some passages in Williamson [140] suggest that he thinks of philosophical exceptionalism as concerning the methods of philosophy as opposed to the methods of science: ‘Of course, philosophy-hating philosophers (a common breed) claim that philosophical thought experiments are profoundly unlike those in natural science, in ways which make the former bad and the latter good, but we should be suspicious of such claims of philosophical exceptionalism’, Williamson [140, p. 179]. Other passages suggest that Williamson thinks of philosophical exceptionalism as something that distinguishes philosophy from ordinary thinking: ‘The main overall aim is to subsume the epistemology of thought experiments under the epistemology of counterfactual conditionals and metaphysical modality (...) and thereby to

characterizes philosophical exceptionalism precisely enough:

Philosophical Exceptionalism: the methods in philosophy are fundamentally different from the methods in the sciences.

The assumption that philosophy has a special kind of method again invites scepticism concerning philosophy more generally.

To sum up, Williamson argues against the *Psychologistic View* because he thinks that it is caused by the false belief in *Evidence Neutrality*. It motivates *Judgment Scepticism* and *Philosophical Exceptionalism*, but leads to scepticism about philosophy more generally and threatens to affect the sciences as well. According to Williamson, the *Psychologistic View*, *Judgment Scepticism*, and *Philosophical Exceptionalism* are all misguided, and we can and should avoid all three by endorsing the *Non-Psychologistic View*.¹²

I will argue that even within a Williamsonian picture of knowledge and evidence, we should neither endorse the *Psychologistic View* nor the *Non-Psychologistic View*. I will suppose that Williamson's inclusive view concerning the evidence we gain from thought experiments (2I) is correct. In general, thought experiments generate knowledge. Hence, in most cases, we know the thought experiment proposition. We also know that we have an intuition towards the thought experiment proposition, and since all knowledge is evidence, both constitute evidence. I will argue that although thought experiments provide evidence consisting of the content of the intuition, it does not follow that we ought to appeal to the content of the intuition when we run thought experiments as counterexamples to philosophical theories. Instead, I think that the psychologistic answer (3P) is correct. I will defend a version of the following view:

reveal it as an application of quite ordinary ways of thinking, not as something peculiarly philosophical.' Williamson [140, p. 180].

¹²In 'Philosophical 'Intuitions' and Scepticism About Judgement' [138], Williamson explains the relations between *Judgment Scepticism* and the *Psychologistic View* slightly differently. He argues that it is common amongst naturalists to doubt our judgments in thought experiments. Being sceptical about judgments in thought experiments, however, is just one instance of being sceptical about judgments in a more general way and endorsing *Judgment Scepticism*. According to *Judgment Scepticism*, thought experiments only provide knowledge of the fact that we have an intuition, which then leads to the *Psychologistic View*.

(2I) The evidence we gain consists in the content of our intuition and in the fact that we have an intuition.

(3P) We ought to appeal *only* to the fact that we have an intuition as evidence.

The version of (3P) I will argue for is meant to be a general methodological rule. In some cases of thought experiments as counterexamples, however, it will be appropriate to appeal directly to the thought experiment proposition as evidence against the target theory. Imagine some undergraduate philosophy student who thinks that knowledge is just belief. In order to convince this student that knowledge is not just belief, we might give a counterexample. Knowledge is obviously not just belief and our counterexample is clearly genuine. In this case, it would not be appropriate to use the fact that we have an intuition and not the content of the intuition as evidence against the theory of knowledge as belief. However, I will defend a version of (3P) for cases of evidence from intuitive counterexamples to theories that are not obviously false. The version of (3P) I will defend is therefore not meant to imply that we should *always* appeal to the fact that we have an intuition rather than to the content of the intuition as evidence from a thought experiment. What I will argue for is that this strategy is always available and the best option in cases where we are genuinely interested in testing a philosophical theory.

In the next section, section 4.4, I defend a version of (3P) by considering Jason Stanley's [116] use of counterexamples to defend his anti-intellectualist account of knowledge.

4.4 Intuitions as Initial Evidence

Stanley introduces his book *Knowledge and Practical Interests* [116] by giving intuitive counterexamples to what he calls 'intellectualism' about knowledge. Intellectualism about knowledge is the thesis that whether someone knows something does not depend upon practical facts, but merely on facts which make the belief more likely to be true, i.e., on truth-conducive facts. According

to Stanley, his cases show that knowledge varies with respect to a non-truth conducive factor, which is the importance the action has to the subject of knowledge.¹³ This, however, is inconsistent with intellectualist theories such as epistemic contextualism. Epistemic contextualism is the semantic thesis that knowledge varies with respect to the context in which it is uttered, where ‘context’ refers to the situation of the subject who is attributing knowledge.

Stanley discusses several variations of two cases originally brought forward by Keith DeRose [26] as evidence for contextualism.¹⁴ Here are Stanley’s versions of the cases:

Low Stakes. Hannah and her wife Sarah are driving home on a Friday afternoon. They plan to stop at the bank on the way home to deposit their paychecks. It is not important that they do so, as they have no impending bills. But as they drive past the bank, they notice that the lines inside are very long, as they often are on Friday afternoons. Realizing that it isn’t very important that their paychecks are deposited right away, Hannah says, ‘I know the bank will be open tomorrow, since I was there just two weeks ago on Saturday morning. So we can deposit our paychecks tomorrow morning.’

High Stakes. Hannah and her wife Sarah are driving home on a Friday afternoon. They plan to stop at the bank on the way home to deposit their paychecks. Since they have an impending bill coming due, and very little in their account, it is very important that they deposit their paychecks by Saturday. Hannah notes that she was at the bank two weeks before on a Saturday morning, and it was open. But, as Sarah points out, banks do change their hours. Hannah says, ‘I guess you’re right. I don’t know that the bank will be open tomorrow.’ [116, pp. 3-4]

¹³Stanley [116, p. 11].

¹⁴DeRose [26, p. 913].

As DeRose [26] has already pointed out, our intuitions are that Hannah is right in both cases. Hannah has knowledge in *Low Stakes* but no knowledge in *High Stakes*. While the contextualist can explain these intuitions, Stanley gives a further case he thinks the contextualist cannot give a straightforward explanation for:

Low Attributor–High Subject Stakes. Hannah and her wife Sarah are driving home on a Friday afternoon. They plan to stop at the bank on the way home to deposit their paychecks. Since they have an impending bill coming due, and very little in their account, it is very important that they deposit their paychecks by Saturday. Two weeks earlier, on a Saturday, Hannah went to the bank, where Jill saw her. Sarah points out to Hannah that banks do change their hours. Hannah utters, ‘That’s a good point. I guess I don’t really know that the bank will be open on Saturday’. Coincidentally, Jill is thinking of going to the bank on Saturday, just for fun, to see if she meets Hannah there. Nothing is at stake for Jill, and she knows nothing of Hannah’s situation. Wondering whether Hannah will be there, Jill utters to a friend, ‘Well, Hannah was at the bank two weeks ago on a Saturday. So she knows the bank will be open on Saturday’. [116, p. 4]

Stanley notes that our intuitive reaction to this case is that Jill is wrong and that Hannah does not know that the bank will be open on Saturday. According to the contextualist, what Jill says would have to be true, because the truth-value of Jill’s knowledge ascription is determined by what is salient to Jill, which is that Hannah has been to the bank two weeks previously, and not Hannah’s current situation, of which Jill is ignorant.¹⁵ According to Stanley’s anti-intellectualism, however, the reason why Hannah knows in *Low Stakes* is that it is not important for her and Sarah to deposit their paychecks, and

¹⁵Stanley [116, p. 24].

the reason why she does not know in *High Stakes* and in *Low Contributor–High Subject Stakes* is that it is important for her and Sarah that they deposit their paychecks. The contextualist, however, cannot give a straightforward explanation of our intuition in *Low Contributor–High Subject Stakes*.

Suppose that, as mentioned above, thought experiments reliably lead to knowledge of the thought experiment proposition, and suppose that most of us have the respective intuitions in Stanley’s cases. Is this all Stanley has to do to refute intellectualism? If Williamson’s view of evidence from thought experiments were correct, Stanley would not need to do anything else but give his counterexamples to intellectualism. However, this does not seem to be all Stanley has to do to refute intellectualism, and in fact he does much more. After presenting his cases, Stanley goes on to discuss alternative explanations of our intuitive reactions to his cases. He rejects the claim that we do not have the respective intuitive reactions, and he considers and rejects the claim that our intuitions are not semantically significant but can be explained by pragmatic aspects of the utterance.¹⁶ If Williamson’s view were correct, there would be no need to give these further arguments. In most of the remainder of his book, Stanley argues against contextualism and other intellectualist theories and for the claim that his view, interest-relative invariantism, does not only get the above cases right but is also the better theory all things considered. It seems that Stanley uses the fact that we have certain intuitions as *initial* evidence, and he then goes on to argue that the truth of the content of these intuitions provides the best explanation as to why we have the intuitions.

By initial evidence, I first of all mean defeasible evidence. Moreover, initial evidence is evidence available given the actual and temporary amount of information or state of investigation. We call a piece of evidence ‘initial’ evidence if more information will be available or if there are more ways to argue that we have not yet considered. Initial evidence might initiate that we refute a theory, but it is not sufficient to refute a theory, and it motivates further investigations.

¹⁶Stanley [116, pp. 12-14].

If we generalize from Stanley’s practice to our general practice of debating intuitive counterexamples to philosophical theories, we get the following version of (3P):

(3P*) We ought to appeal *only* to the fact that we have an intuition, and we ought to appeal to this fact as *initial* evidence only.

To support (3P*), I present some strategies which are used in our practice of debating intuitive counterexamples in philosophy in the next section.

4.5 Intuitive Counterexamples

I will first present some strategies we use to show that a counterexample is merely intuitive as opposed to genuine. The strategies are not intended to be exhaustive, but they provide a useful categorization of our practices. The examples I use might be controversial, i.e., different philosophers might have different views on the correct way of describing a particular case and which strategy is appropriate for it. However, the fact that we commonly use these strategies should not be controversial.

The first strategy is to deny the relevance of certain intuitions due to some psychological disposition of the person having the intuitions. The following three are strategies of explaining away the significance of a particular intuition. By ‘explaining away the significance of an intuition’, I mean explaining why we have the intuition even though the content of the intuition is not true, which involves showing that the intuition is not significant to the truth or falsity of its content.¹⁷

¹⁷See also Ichikawa [52]. Such an explanation can rely on some pragmatic rules of conversation, on some properties of our reaction to the presentation of the thought experiment scenario or other irrelevant psychological features, or on some properties of the theory’s subject matter. They all presuppose that we use the fact that we have an intuition as evidence. The fourth strategy involves re-identifying the content of the intuition or the target theory, which shows that we use the evidence we gain from thought experiments as initial evidence only.

(i) Relevance of the intuitions

Experimental philosophers have shown that laypeople's intuitions in thought experiments are subject to all kinds of irrelevant factors, and they think it is likely that philosophers are prone to the same effects, which might influence and often mislead their intuitions. I mentioned in Chapter 3 that philosophers such as Kornblith [65], Ludwig [78], and Williamson [140] think that the experimental philosophers' studies are irrelevant because we are interested in expert intuitions only rather than in laypeople's intuitions. They take it that philosophers' intuitions in thought experiments are more reliable than laypeople's intuitions because like experts in other disciplines, philosophers have specific skills and specific philosophical knowledge and experience which makes them better intuiters.¹⁸ Kornblith claims that like in the sciences, where 'accurate theory is allowed to play a role in guiding and shaping observation', we should not prefer 'the intuitions of the ignorant and the naive over those of responsible and well-informed investigators'.¹⁹

Experimental philosophers disagree about which intuitions are relevant. According to Adam Feltz & Chris Zarpentine [35] and Ram Neta & Mark Pheelan [88], empirical evidence has cast doubt on the evidence Stanley uses to defend anti-intellectualism. The authors claim that Stanley's defence of anti-intellectualism is based on the empirical assumption that it captures part of our ordinary practices of knowledge ascription. Feltz & Zarpentine write:

[...] Stanley's case for anti-intellectualism depends on 'examples that suggest' that ordinary knowledge ascriptions are sensitive to the practical facts of the subject's situation (2005, p. 3). But are our ordinary knowledge ascriptions sensitive to the practical facts of a subject's situation? [We] investigate this question empirically
 [...] [We] focus on Stanley because his case against intellectual-

¹⁸Ludwig [78] compares philosophers to mathematicians, Williamson [139], [140] to lawyers, and Weatherston [132] to business people.

¹⁹Kornblith [65, pp. 33-34].

ism clearly depends on facts about ordinary knowledge ascriptions. Thus, we treat Stanley's claims about 'the intuitive reactions we have to these cases' (2005, p. 5) as predictions about ordinary knowledge ascriptions. Our results do not support Stanley's predictions. If our data generalize, the case for anti-intellectualism cannot depend on the sensitivity of our ordinary practices of knowledge ascription to the practical facts of a subject's situation. Ordinary knowledge ascriptions do not appear to be sensitive to these facts. [35, p. 6]

Feltz & Zarpentine's empirical studies on 152 students in introductory-level philosophy classes suggest that anti-intellectualism does not generalize to people's ordinary practice of ascribing knowledge in some paradigmatic cases. People's intuitions in these paradigmatic cases do not show the pattern that would support anti-intellectualism. Hence, the intuitions Stanley and others might have in Stanley's cases are not relevant to the question whether intellectualism is true.

In line with Kornblith, Ludwig, and Williamson, Stanley could argue that the fact that these students have certain intuitions is not relevant with respect to whether his cases support anti-intellectualism, e.g., because they are not experts (see the discussion of expertise in philosophy in section 3.6 of Chapter 3).

(ii) Explaining away: pragmatic explanations

Explaining away the significance of an intuition to the truth or falsity of the proposition in question by relying on some pragmatic rules of conversation has been a very common practice in philosophy of language since H. P. Grice's [50] introduction of conversational implicature. It is widely accepted that our use of language is not only guided by semantic, but also by pragmatic rules. There is a difference between the information we usually communicate with a

sentence and the meaning of the sentence, and what we usually communicate affects some of our intuitions. When we search for semantic facts, we therefore have to make sure that we do not confuse them with merely pragmatic ones.

Here are three examples provided by Kent Bach [3]:

1. Jack and Jill went up the hill.
2. Jack and Jill are engaged.
3. Jill got married and became pregnant.

According to Bach, our intuitions about what these sentences mean are likely to be biased in favour of understandings corresponding what people are likely to communicate. In uttering sentences (1) - (3), a speaker is likely to mean something more specific, namely the following:

- 1*. Jack and Jill went up the hill together.
- 2*. Jack and Jill are engaged to each other.
- 3*. Jill got married and then became pregnant.

Since we take (1*) - (3*) to be most typically communicated, we intuitively take it to be the meaning of the sentences (1) - (3). However, what we are likely to communicate with these sentences is not necessarily their meaning. These three examples might be very easy to detect and one might think that philosophers are not prone to a semantic/pragmatic confusion. However, there are much more complicated cases where such a confusion is not as easy to detect. Bach [4] gives a variety of cases, some of which seemed to have been rather difficult to figure out:

Overlooking the semantic/pragmatic distinction led the so-called ordinary language philosophers to make some misguided claims about the meanings of such philosophically interesting terms as 'tries', 'seems', 'good', and 'true'. They were right to ask what we would mean in uttering sentences containing such terms, but they

were wrong to draw pragmatics-laden conclusions about the meanings of these terms. For example, it was thought that trying to do something entails some effort or difficulty in doing it and that something's seeming to have a certain property entails doubt if not denial that it does have that property. These alleged entailments actually stem from conditions for appropriately and non-misleadingly using sentences containing 'try' or 'seems', not from their truth-conditions.

Bach thinks that it is questionable to what extent our intuitions in general actually reveal semantic facts, i.e., to which extent they should be explained by a theory rather than just explained away, hence they 'should be relied upon judiciously'.²⁰ Keith DeRose [27], in contrast, warns that pragmatic explanation should not be exaggerated if we want the method of counterexamples to be a useful practice, and Jason Stanley [116] thinks that pragmatic explanations have to involve general conversational principles such as Gricean maxims of conversation in order to provide a satisfactory alternative to taking the intuition at face value.²¹ Whereas there is disagreement about the extent to which we should apply this practice and what the requirements are, there is no doubt about the danger of confusing pragmatic facts with semantic facts in our intuitions.

(iii) Explaining away: psychological explanations

The strategy of empirically testing and explaining why we have a certain intuition on the basis of psychological theories has recently become popular in philosophy, and the relevance of empirical work to traditional philosophical questions has been extensively discussed in the recent literature on the methodology of philosophy.

²⁰Bach [3], see also [4].

²¹Stanley [116, pp. 13-15].

It seems that experimental philosophers have further developed a strategy which has, until recently, been pursued from the armchair. Philosophers have always tried to vary some factors of the thought experiment scenario in order to rule out alternative explanations as to why we have a certain intuition. Here are two quotes from Stanley [116] on his cases presented above, where he argues against two alternative psychological explanations of our intuitions:

For example, one might attempt to explain away the force of the intuitions behind these scenarios, by arguing that, when someone recognizes that the cost of being wrong are particularly high, his or her confidence is shaken. [116, p. 6]

Another strategy that proponents of intellectualism commonly appeal to in the face of these examples is to argue that in certain cases our responses are sensitive not to whether the subject knows, but to whether the subject *knows* that she knows. [116, p. 7]

It is obvious that both of these claims are psychological theses. As long as we have not tested these theses, that our confidence is shaken and that we confuse knowledge with knowledge of knowledge in the respective cases are not much more than speculations about how our mind works, and running experiments or backing them up with psychological evidence of some kind in order to evaluate them would clearly be an improvement.

(iv) Explaining away: ontological explanations

We can also explain away the relevance of an intuition by relying on a property of the theory's subject matter. Brian Weatherson [131] gives four criteria by which we can judge a philosophical theory. The first criterion is that there should not be too many intuitive counterexamples to it. According to Weatherson, intuitions come in various degrees, and whether a counterexample is a threat to the theory depends on how strong the intuition is. Also, the more 'obscure or fantastic' a counterexample is (involving zombies, etc.), the less

damaging it is to a theory.²² The second feature a theory should have is that there should not be too many unacceptable theoretical consequences. Such unacceptable consequences can, for instance, be caused if we incorporate too many intuitions into the theory. A theory should moreover pick out significant properties of the subject matter in question and should also be as simple as possible. These four criteria can be in conflict with each other and have to be weighted against each other. As a consequence, a theory can be true if there are intuitive counterexamples to it.

In order to explain why an intuitive counterexample might cause no harm to a philosophical theory, Weatherson endorses a Lewisian theory of meaning. When we assign a property to a predicate, we must make sure that first, this property satisfy as many as possible of our pre-theoretic beliefs, and secondly, that it is simple and theoretically important. According to Lewis, for a property to be simple and theoretically important is to possess a primitive naturalness. However, our pre-theoretic beliefs sometimes do not cover the naturalness of the property. According to Weatherson, this might be the case with respect to the JTB theory of knowledge: there are no reasonably natural properties in the neighbourhood of our disposition to use ‘knows’. If this is right, then even some significant changes to our use are not changes in meaning of ‘knows’, because they do not change which is the closest reasonably natural property to our usage pattern of ‘knows’. Hence, the JTB theory as the best trade-off ‘is still a live possibility, even considering Gettier Cases’.²³ Gettier gave an intuitive counterexample, but the JTB theory does not have many unacceptable theoretical consequences, it is significant and extremely simple. In other words, the JTB theory does very well on the three remaining criteria, so that the Gettier Cases might not cause any harm to it. Hence, a theory as strong as the JTB theory of knowledge itself provides good evidence against the genuineness of the intuitive counterexample.

The practice of giving an explanation that relies on a property of the subject

²²Weatherson [132, p. 8].

²³Weatherson [131, p. 11].

matter is probably least common. It is quite different from the practice of explaining away the significance of an intuition by appealing to pragmatic rules or by giving a psychological explanation as to why we have the intuition. One reason is that it presupposes a controversial theory of meaning. Moreover, while pragmatic or psychological explanations can predict our intuitions on the basis of some general conversational rule or some underlying psychological feature, giving an explanation that relies on a property of the subject matter only cannot explain or predict our intuitions.

(v) Identifying the content of the intuition

After having given a thought experiment as a counterexample, philosophers sometimes discover that the thought experiment proposition actually is not inconsistent with the target theory. I will give two examples.

Vann McGee [84] argues that *modus ponens* is not an entirely reliable rule of inference. According to McGee, there are occasions in which one has good reasons to believe the premises but no reason to believe the conclusion of an argument. He gives, among others, the following case:

Having learned that gold and silver were both once mined in his region, Uncle Otto has dug a mine in his backyard. Unfortunately, it is virtually certain that he will find neither gold nor silver, and it is entirely certain that he will find nothing else of value. There is ample reason to believe

If Uncle Otto doesn't find gold, then if he strikes it rich,
it will be by finding silver.

Uncle Otto won't find gold.

Since, however, his chances of finding gold, though slim, are no slimmer than his chances of finding silver, there is no reason to suppose that

If Uncle Otto strikes it rich, it will be by finding silver.

[84, p. 463]

McGee takes this case to be a counterexample to the general validity of *modus ponens*. However, other philosophers have replied that it merely shows that *modus ponens* does not necessarily provide reasons to believe the conclusion, given that one believes the premisses. It does not show that the inference rule is not valid. If this is true, it means that McGee's case is consistent with and hence not a genuine counterexample to *modus ponens*.²⁴

Some philosophers have argued that Frank Jackson's Mary Case against physicalism does not show that physicalism is false, but that it reveals interesting facts about other kinds of knowledge. Here is Jackson's case:

Mary is a brilliant scientist who is, for whatever reason, forced to investigate the world from a black and white room via a black and white television monitor. She specialises in the neurophysiology of vision and acquires, let us suppose, all the physical information there is to obtain about what goes on when we see ripe tomatoes, or the sky, and use terms like 'red', 'blue', and so on [...] What will happen when Mary is released from her black and white room or is given a colour television monitor? Will she learn anything or not? [54, p. 130]

Jackson's original answer was that it seems 'obvious that she will learn something about the world and our visual experience of it.'²⁵ He concluded that Mary's knowledge about the neurophysiology of vision must have been incomplete even though she had all the physical information. Hence, there is more knowledge to have than knowledge about the physical and physicalism is false.²⁶

²⁴For a critical discussion, see Sinnott-Armstrong, Moor, and Fogelin [105], Lowe [77] and Nolan [90]. Lycan [80] accepts the case as a genuine counterexample to *modus ponens*.

²⁵Jackson [54, p. 130].

²⁶Jackson has since changed his view and agrees with his critics that the Mary Case is not a genuine counterexample to physicalism (e.g., [55]).

The problem is that it is not obvious what exactly the content of the intuition that Mary learns something is. Here are two alternative suggestions, according to both of which the case shows something about our knowledge but is not inconsistent with physicalism. The first has been defended by David Lewis [74], [75]. According to Lewis, what Mary acquires is knowledge-how as opposed to knowledge-that. She learns how to recognize, remember or imagine colours, e.g., how to recognize a red tomato or to remember or imagine the blue sky. Learning how to do these things does not amount to learning facts, it is to acquire a bundle of abilities. Therefore, what Mary learns when she leaves the black and white environment is not incompatible with the completeness of her factual knowledge of the physical information about vision. According to Earl Conee [18] and Michael Tye [123], Mary neither acquires knowledge-that nor knowledge-how. What she acquires is knowledge by acquaintance in the sense that she becomes familiar with redness in the most direct possible way. Whereas Mary knows everything about the properties of experiences, she gets acquainted with them only when she leaves her black and white environment. Both views just presented have it that Mary knows everything about the physical facts and hence physicalism is true, but she acquires *non-propositional knowledge* when she leaves her black and white environment.

In both cases, the intuitive counterexample turns out not to be genuine, because the content of the intuition is not in conflict with the original target theory.

I presented five strategies of showing that an intuitive counterexample to a philosophical theory is not genuine. Even though the extent to which these strategies ought to be applied might be controversial, they are commonly used in contemporary analytic philosophy. They show that we often do more than just using the thought experiment proposition in order to refute the target theory.

Strategy i) suggests that we sometimes question the relevance of certain

intuitions, in particular of intuitions of a certain group of people. Strategies ii), iii), and iv) suggest that even though our intuitions from thought experiments are reliable, there are several ways to question the truth tracking properties of a particular intuition. Strategy v) shows that there are ways to question whether the thought experiment creates an inconsistency with the target theory. This strategy suggests that we sometimes need more evidence as to whether the thought experiment proposition is in fact inconsistent with the target theory. It seems that all strategies are best explained by our use of the fact that we have an intuition as initial evidence. Had we used the content of the intuition as evidence, different ways of explaining why we have the intuition would not play any role.

There is a further strategy which we use when we consider thought experiments as counterexamples to philosophical theories, which is to provide independent arguments in support of the content of the intuition. Max Deutsch [29] thinks that there is no widespread practice of appealing to the fact that we have an intuition as evidence against a philosophical theory. One of his arguments for this claim is that there is not *one single* way to reject a philosophical theory. While I argued in Chapter 3 that he does not succeed in establishing that intuitions play no role as evidence, I agree that we often use additional arguments to support the content of the thought experiment proposition. This, I think, supports my claim that we use the evidence we appeal to from a thought experiment as *initial* evidence only.

As an example, Deutsch shows that Gettier and others have provided multiple arguments against the JTB theory of knowledge. Here is a quote from Deutsch's paper, where he gives several examples:

- (a) [...] S's justified true belief that p might fail to be knowledge if there is a disconnection between: (i) what causes S to believe p, and (ii) what makes S's belief that p true. In the 10 coins-case, it is the number of coins in Jones's pocket that is (partly) causally responsible for Smith's belief that the man who will get

the job has 10 coins in his pocket. But that belief is true ‘in virtue of’ the number of coins in Smith’s pocket. According to Gettier, this disconnection between what causes Smith to believe and what makes Smith’s belief true justifies the judgment that Smith does not know.

(b) Some epistemologists (e.g. Unger (1968), [...] Pritchard (2005)) argue for the related idea that agents in (at least some) Gettier Cases are only luckily correct in believing what they do, and that the presence of this sort of ‘epistemic luck’ explains why such agents lack knowledge. In the 10-coins case, it is an accident (luck) that Smith is right in believing that the man who will get the job has 10 coins in his pocket, given that this belief is based partly on a count of the coins in the pocket of a man who, as it happens, is not the man who will get the job. This appears to be compelling grounds for concluding that Smith’s belief does not add up to knowledge.

(c) Other epistemologists (e.g. Lehrer and Paxson (1969)) claim that S’s justified true belief that p may fail to qualify as knowledge if there are ‘epistemic defeaters’ to S’s justification. Roughly, q is defeater to S’s justification for p, if q is compelling evidence against p about which S is unaware. The presence of defeaters might explain Smith’s lack of knowledge in the 10 coins-case. If Smith were apprised of the fact that it was he who would get the job, not Jones, then Smith would no longer justifiably believe that the man who will get the job has 10 coins in his pocket. Smith’s justification for his belief is thus defeated and, for this reason, arguably fails to count as knowledge. Deutsch [29, p. 14]²⁷

To summarize, Deutsch first points to Gettier’s argument that the *disconnection* between what causes Smith to believe and what makes Smith’s belief

²⁷References: Goldman [44, pp. 357-372]; Unger [124, pp. 157-170]; Pritchard [94]; Lehrer & Paxson [72, pp. 225-237].

true justifies the judgment that Smith does not know. He then points to the fact that other epistemologists have argued that the *luck* involved in the person's belief and the truth of this belief make the case such that the person does not have knowledge. Finally, he draws attention to the argument that unrecognized defeaters prevent the person from having knowledge. All these arguments have been given independently from the intuition we have in the Gettier Cases in support of the counterexample.

I have presented several strategies we use to either show that an intuitive counterexample is merely intuitive or to show that it is genuine. Some of these strategies presuppose that we use the fact that we have an intuition as evidence. However, that we use these strategies only shows that (3P*) is a common practice, rather than that it is a practice we ought to pursue. One might want to object that while (3NP) should be our standard practice, we step back and appeal to the fact that we have an intuition only when challenged. Hence, only upon challenge, our practice ought to be (3P*).

In Chapter 5, I will give a more general defence of (3P*). As a short reply, let me say the following. First, given that there are different ways of explaining the fact that we have an intuition which are commonly used, it would be inappropriate not to consider them when considering an intuitive counterexample. Second, one might say that in the case of an intuitive counterexample to an accepted philosophical theory, we are always in a situation where our evidence is challenged, because the target theory is in conflict with the thought experiment.

In the following two sections, I address two worries Williamson raises with respect to (3P). I show that (3P*) does not provoke a regress of giving evidence for evidence (section 4.6) and does not presuppose or lead to *Philosophical Exceptionalism* (section 4.7).

4.6 The Regress Argument

Let us have a closer look at some of Williamson's arguments against a view which entails (3P), the claim that we ought to use the fact that we have an intuition as evidence from thought experiments. As mentioned above, when we run a thought experiment, we usually do not only gain knowledge of the thought experiment proposition, we also gain knowledge of the psychological fact that we have an intuition that the thought experiment proposition holds. Because all knowledge is evidence, this piece of knowledge could be appealed to as evidence as well. Williamson [138] considers this option. He says that someone might respond to his view entailing (3NP) thus:

Granted, when we are consciously inclined to judge that P, we often but not always know that P. That we are consciously inclined to judge that P should therefore be treated as good but defeasible evidence for the claim that P. It is just one more part of the total body of evidence on which philosophical theories should be evaluated. [138, p. 200]

However, Williamson thinks that appealing to only this part of our evidence would not be a good practice, for the following reason:

What [this reply] perversely ignores is the evidential role of the fact that P itself, as opposed to that of the fact that we are consciously inclined to judge that P. After all, if we do know that P, would it not be negligent not to use that knowledge in evaluating a philosophical theory to which it is relevant? [138, p. 148]

According to Williamson, one might argue that if we do not know whether we know the thought experiment proposition, we cannot use our knowledge to evaluate the target theory. What we need is knowledge of knowledge of the thought experiment proposition. Williamson argues that this view is inadequate for two reasons:

First, it gives no more reason to deny that we know that we know that P than to deny that we know that P in the relevant cases. Although we cannot expect to have infinitely many iterations of knowledge, for more than computational reasons (Williamson 2000, 114-34), that general point merely shows that we must sometimes simply apply our knowledge, without first checking whether we know, for otherwise we get stuck in an infinite regress of checks. That is the second problem for the envisaged defence of [the above reply]. [138, p. 149]

While I will not argue that we need knowledge of knowledge of the thought experiment proposition, I think it is correct that we should use the fact that we have an intuition as evidence from thought experiments as initial evidence and consider alternative explanations as to why we have the intuition. Williamson's worry is that we might end up in an infinite regress of giving evidence for evidence. He concedes, however, that this is not a practical concern: at a certain point, we just apply our knowledge. The question is whether there is an epistemic reason to stop it.

Let us suppose that the fact that we have an intuition in a thought experiment can be used as initial evidence against the target theory, as $(3P^*)$ has it and as our practice suggests. There is a possible regress at two different points. First, we might be worried about evidence for the fact that we have an intuition that P . Second, we might be worried that we need evidence for the further evidence we use to show that the counterexample is genuine.

As *initial* evidence, the fact that we have an intuition is evidence given the momentary state of information and investigation, and further evidence is required to show that the counterexample is genuine and actually refutes the target theory (as I have argued in sections 4.4 and 4.5). Now, if we investigate further and eventually come to the conclusion that the counterexample is genuine, i.e., does in fact undermine the target theory, it does not matter whether the counterexample is intuitive or not. Not every counterexample is intuitive,

and since what counts against the theory is whether the counterexample is genuine, it does not matter whether we really find it intuitive or it just seems to us that it is intuitive. If, however, upon further investigation, we come to the conclusion that our counterexample is not genuine, we do not need to know whether it is intuitive or not either. Hence, from an epistemic point of view, it is redundant to investigate whether we actually have an intuition that P or not, because this is not what we are ultimately interested in. If we more often than not actually have an intuition that P when it seems to us that we have an intuition that P in a thought experiment, investigating whether an intuitive counterexample is genuine is the strategy required. It seems that there are not just practical reasons to stop the regress; there is, from an epistemic point of view, a redundancy to giving evidence for the claim that we have discovered an intuitive counterexample.

The second regress one might fear concerns the question of when we are allowed to stop further investigations into the truth or falsity of the thought experiment proposition. How can we determine whether we have shown that a thought experiment undermines a philosophical theory? Do we have to run through all possible strategies mentioned in section 4.4? This seems to be a question that concerns philosophy in general and to which I cannot give an answer. The sceptic who challenges us here would have to be a sceptic about arguments and explanations. The sceptic would have to endorse something as general as the following:

Scepticism about Arguments and Explanations: we cannot trust our arguments and explanations in philosophy.

It will be very hard to endorse such a general version of scepticism about philosophy without it affecting the sciences as well.

In this section, I have argued that following (3P*) as a general methodological rule is a good practice because our intuitions only serve as motivation for further investigations into the truth and falsity of the thought experiment

proposition, and using the fact that we have them as initial evidence does not lead to a regress of giving evidence for evidence. We have epistemic reasons to stop the regress. I have argued that the psychological facts are not the only evidence and that we also use evidence from arguments, etc. Since the psychological facts are only initial evidence, there is less need to address the problem of the gap between facts about our psychological states and facts about the world. The reason why we use the fact that we have an intuition as initial evidence only does not seem to be that we do not trust our judgments, it is rather that refuting a philosophical theory involves more than just knowledge of the thought experiment proposition.

In the next section, I show that scientists proceed similarly when they discover a potential counterexample to an established scientific theory.

4.7 Philosophical Exceptionalism

Williamson thinks that our evidence in philosophy does not consist in facts with respect to which we are infallible: we sometimes might not have knowledge in the case of a thought experiment. However, successful counterexamples such as the Gettier Cases suggest a general strategy as to what kind of knowledge we need in order to refute a philosophical theory:

[ordinary] knowledge is enough. We have no general guarantee against the possibility that we did not know something that we thought we knew. [138, p. 151]

This, Williamson thinks, is just parallel to what we do in the empirical sciences.

I will suggest that when scientists discover a counterexample to an established theory, they do not simply apply their knowledge, even if the methods they applied to gain that knowledge are reliable. If this observation is correct, taking the fact that we have an intuition as initial evidence is not in principle different from what scientists do and hence does not presuppose or lead to *Philosophical Exceptionalism*.

Here is a recent example where scientists discovered evidence against an established theory. In spring 2009, concern was raised in Canada that prior vaccination with the common influenza vaccine was associated with an increased risk of swine-flu (pH1N1 virus). This concern led to several epidemiologic investigations in order to assess the putative association. The conclusions of the studies were that prior vaccination with 2008-09's seasonal influenza vaccine was associated with an increased risk of swine-flu during spring and summer 2009 in Canada. Here is a quote from the introduction to the study:

In this paper we report the expected finding that 2008-09 TIV was associated with a significant [...] reduction in the risk of medically attended illness due to seasonal influenza. However, we also report the unexpected finding that TIV receipt was subsequently associated with a statistically significant [...] increased risk of medically attended illness due to the novel pH1N1 virus [...]. Because the latter result is contrary to established knowledge, greater scrutiny is required to determine whether these associations are more likely on balance to be real (causal) or due to a methodological flaw (bias).
[106, p. 8]

It is important to note that these results are unexpected. First, they contradict the established theory. Second, some other studies conducted in other countries have shown that seasonal vaccination 'had no influence or may have been associated with reduced chances of pH1N1 illness'. The scientists have data that suggest that the seasonal flu vaccine may increase a risk of getting infected with swine-flu, but further investigations are required in order to confirm these data and possibly eventually reject the theory that there is no association between the seasonal vaccination and the risk of getting infected with swine-flu. The data can certainly be used as initial evidence for the fact that there is such an association, and hence as initial evidence against the established theory.

Scientists have different strategies to find out whether their data correspond to the facts in the world. Methods based on observation are susceptible to random variation, which means that the vaccinated people who took part in the study might show the discovered feature only by chance. Furthermore, there could be some bias in the selection of people or in the information that was provided by the people who took part in the study. It might also be that the vaccinated and tested people share another unknown characteristic that is actually responsible for increasing their risk of developing swine-flu.

My suggestion is to view the method of thought experiments as counterexamples to philosophical theories as similar to conducting experiments that produce counter-evidence to established scientific theories. Obviously, there are differences between philosophy and the sciences when it comes to the methods. However, these differences do not affect our more general practice of dealing with counter-evidence to established theories. Even if it is likely that the scientists know that there is an association between the seasonal vaccination and the risk of getting infected with swine-flu because they applied a reliable method, they need to exclude other explanations of their data in order to refute the established theory. A similar claim is true for philosophy. Even if it is very likely that we know the thought experiment proposition, we need to know whether the counterexample is genuine as opposed to merely intuitive, which involves more than knowing the thought experiment proposition.

Williamson says the following about thought experiments as counterexamples to philosophical theories:

Philosophy is hard enough already: why should we make it even more difficult by forbidding ourselves to bring some of our knowledge to bear? We are not obliged to fight with one arm tied behind our back. [138, pp. 148-149]

Without question, science is hard as well. However, it is obvious that we would not want to say something similar in the sciences: ‘Science is hard enough

already: why should we make it even more difficult by forbidding ourselves to bring some of our knowledge to bear? We are not obliged to fight with one arm tied behind our back.’ Because of what is at risk, it is easy to see that mere knowledge is not enough when it comes to counter-evidence to an established scientific theory. We need to know whether we know that there is an association between the seasonal vaccination and the risk of getting infected with swine-flu. I think that the same holds for philosophy: in order to have refuted the target theory, we need to know that the thought experiment is a genuine counterexample that undermines the target theory.

4.8 Conclusion

The aim of this chapter was to argue that a view of evidence from thought experiments combining (2I) and (3P*) corresponds with our practice of debating intuitive counterexamples to philosophical theories. While the evidence we gain consists in the content of the intuition and in the fact that we have an intuition (2I), we appeal to the fact that we have an intuition as initial evidence in order to refute a philosophical theory, which corresponds with (3P*).

I showed in sections 4.4 and 4.5 that there are various strategies we use in philosophy which presuppose (3P*). I suggested that our practice shows that we use the fact that we have an intuition in a thought experiment only as initial evidence against a philosophical theory, and that the main function of our intuitions is to motivate further investigations as to whether the counterexample is genuine.

I argued in sections 4.6 and 4.7 that Williamson’s regress worries do not apply if the fact that we have an intuition is used as initial evidence only and that the practice of appealing to the fact that we have an intuition as initial evidence does not lead to or presuppose *Philosophical Exceptionalism*: when it comes to counter-evidence to an established theory, our practice in philosophy is not in principle different from our practice in the sciences.

So far, I have not given an explanation as to why appealing to the psychological fact that we have an intuition as evidence from thought experiments is a good practice. I address this issue in the next chapter, Chapter 5.

Chapter 5

Knowledge and Evidence in Philosophy

Abstract

In Chapter 4, I argued that actual philosophical practice standardly involves appeal to our intuitions rather than to the contents of our intuitions. In this chapter I offer an explanation of why this is the case. I first show that there are methodological constraints on evidence in the sciences, which suggests that there are such constraints in philosophy as well. I discuss several options of what the constraints on evidence in philosophy could be. I then defend an account according to which we ought to appeal to the part of our total evidence which is relatively easy to access.

5.1 Introduction

Supposing that thought experiments reliably lead to knowledge of the content of our intuition that P , we can distinguish the following two views concerning the evidence we usually gain from a thought experiment (where P is true):

Psychologistic View: the evidence we gain from a thought experiment consists in the fact that we have an intuition that P .

Inclusive View: the evidence we gain from a thought experiment consists in the fact that we have an intuition that P and in the fact that P .

In Chapter 4, I distinguished between the evidence we gain from a thought experiment and the evidence we ought to appeal to when we aim to refute a philosophical theory. I argued that using the fact that we have an intuition that P as *initial* evidence from thought experiments corresponds with our practice in philosophy. In this chapter, I further motivate the claim that this is a practice we ought to pursue even if we endorse the *Inclusive View*.

In the first part (sections 5.2 and 5.3), I present cases which suggest that there are methodological constraints on evidence in the sciences (section 5.2), and I defend this view against alternative interpretations of the cases (section 5.3). In the second part (section 5.4), I argue that there are methodological constraints on evidence in philosophy as well, because philosophy is similar to the sciences with respect to some relevant aspects, which makes it plausible that there are such constraints on evidence in philosophy as well. Timothy Williamson [138], [140] offers two explanations as to why many contemporary philosophers appeal to the fact that they have an intuition that P . In ‘Philosophical ‘Intuitions’ and Scepticism About Judgement’ [138], Williamson claims that some might believe evidence has to meet an operational standard, and they think that the fact that someone has an intuition that P meets this standard, but the content P does not. In *The Philosophy of Philosophy* [140, pp. 208-215], he argues that many contemporary philosophers appeal to the fact that they have an intuition as evidence because they falsely believe in *Evidence Neutrality*, which is the idea that a community of philosophers can always in principle achieve common knowledge as to whether any given proposition constitutes evidence or not. I consider an operational standard and *Evidence*

Neutrality as methodological norms as well as *Evidence Neutrality* as a conversational norm. In the third part (section 5.5), I then suggest a methodological norm for evidence in philosophy according to which we ought to appeal to the part of our total evidence for a certain claim which is relatively easy to access. This norm can explain why we appeal to the fact that we have an intuition that *P* as evidence from a thought experiment even if we endorse the *Inclusive View*, and I suggest that it explains our use of evidence in many other cases as well.

5.2 Evidence in the Sciences

Not everything we take to be evidence in ordinary life contexts can be used as evidence in the sciences. Scientific evidence has to meet standards such as statistical relevance, objectivity, and reproducibility. Moreover, the methods by which we acquire scientific evidence have to be made transparent. People not familiar with scientific practices do not always seem to appreciate this fact, as the following shows.

In an online article titled ‘Duh! The Most Obvious Scientific Findings of 2010’¹, two authors list scientific findings such as ‘Caffeine affects kids’ sleep’, ‘Smoking a lot of weed is bad for you’, or ‘Sitting in front of the TV all day can make a teen fat’. They introduce their article with the line ‘Along with some truly groundbreaking discoveries, scientists this year told us a few things we already kind of knew’. Under the headline ‘Bullies pick on unpopular kids’, they ridicule a study by René Veenstra et al. [127] published in *Child Development*. They write ‘Who’d have guessed? Bullies target kids who are unpopular and less likely to be defended by their peers, a new study finds.’ Several other online newspapers and blogs make fun of the study, some of them calling it a waste of money.

Here is how Veenstra et al. gained their evidence. For their study on bul-

¹<http://www.livescience.com/11227-duh-obvious-scientific-findings-2010.html>. Date: 29 December 2010.

lying, they tested 26 elementary school classes. The research assistants first explained to the children what bullying is.² They then gave examples of bullying and of forms of behaviour that should not count as bullying (e.g., teasing in a friendly and playful way). The children finally had to fill out questionnaires where they answered questions concerning which peers they accepted or rejected, and who they were bullying or bullied by.

The results of a study by Britt Hedman Ahlström [1], titled ‘Major Depression and Family Life—The family’s way of living with a long-term illness’, was listed amongst the most obvious findings of 2009. The authors of the online article comment on it as follows:

This sad truth stands to reason: Life is hard for kids whose parents are depressed. ‘Children take on an extremely heavy responsibility by monitoring and keeping an eye on the depressed parent,’ said study author Britt Hedman Ahlström of the Sahlgrenska Academy at the University of Gothenburg, Sweden. Hedman Ahlström found that a parent’s depression increases a child’s sense of responsibility and feelings of loneliness. She recommends that health services support the whole family, and not just the person who is ill [...] *Good advice, and hopefully not news to the mental health community.*³

Let us see how Hedman Ahlström conducted her study on depression and family life. Hedman Ahlström conducted 25 interviews, both group interviews and individual interviews, which were all tape-recorded and transcribed. She used qualitative content analysis and phenomenological hermeneutic method to analyze the data. Qualitative content analysis requires the researcher to divide the interview text into units, sample the units in accordance with the research question and rewrite it in categories and themes.⁴ Phenomenological hermeneutics is performed in three steps: the researcher reads

²They used the term ‘bullying’ as defined in the Olweus’ [91] Bully/Victim questionnaire.

³<http://www.livescience.com/5965-duh-year-obvious-discoveries.html>, emphasis added. Date: 31 December 2009.

⁴Hedman Ahlström [1, p. 20]. Hedman Ahlström refers to Krippendorff [67], amongst others.

the text naively, then structures and analyses it and finally reaches an interpretation on the basis of the first two steps and of independent knowledge.⁵

The strong reactions to seemingly unnecessary studies quoted above can be explained by the fact that many people are not familiar with scientific methods and constraints on scientific evidence. To scientists, however, it should not come as a surprise that Veenstra et al. and Hedman Ahlström tested their hypotheses empirically, using scientifically recognized methods such as questionnaires and interviews, as well as certain methods to analyze the data gained through these questionnaires and interviews. On the basis of these two cases and how they were discussed, we can construct the following fictional cases.

Sally the Teacher

Sally is a teacher. She has been teaching for a couple of years and has been observing a lot of bullying in her classroom. Sally comes to believe that bullies choose victims that are already rejected by their classmates, which gives them a rise in status on the one hand and no loss in affection by other classmates on the other hand. She speaks to many teachers in her school who have observed similar situations and confirm Sally's observations about bullying. Given her evidence, Sally knows that bullies choose victims that are already rejected by their classmates, and she acts appropriately when she takes precautions to protect potential victims from potential bullies. After some more years of teaching, Sally goes back to university and carries out a study that confirms exactly what she already knew: bullies want to realize both status and affection and therefore choose victims that are already disliked by their classmates.

⁵Hedman Ahlström [1, p. 22]. The method of phenomenological hermeneutic was developed, amongst others, by Lindseth & Norberg [76].

There seems to be nothing wrong with this case. Before the study, Sally knows that bullies choose victims that are already disliked by their classmates, and that they do so to realize both status and affection. The propositions are true, Sally believes them and she has enough evidence through her own experience and through testimony. Sally also appropriately takes precautions to protect potential victims from potential bullies. From a *scientific* point of view, however, running a study is necessary. Imagine, for instance, the psychology department is planning to edit a new textbook on education, and they ask Sally to write a contribution on bullying. It would be unacceptable for Sally to publish her knowledge gained in the classroom and through testimony (e.g., anecdotally) in the textbook. In order to meet scientific standards of evidence, she clearly needs to run the study.

Here is a second case:

Sam the Psychologist

Sam is a psychologist who works with patients suffering from major depressions. Many of Sam's patients are parents of children, and many of them raise concerns with regard to their children's mental health. Sam decides to talk to some of the children as well, and over the years, she realizes that their mental health is either already heavily affected by their parent's disease or is threatened to be affected. Sam decides to change her practice: from now on, she insists on treating the children as well whenever a parent suffers from a major depression. Given her evidence, she knows that a major depression of a parent negatively affects the children's mental health, and she acts appropriately when she insists on treating them as well. After a couple of more years, Sam is up for a new challenge. She decides to run a study on her patients and publish the results. Her study confirms exactly what she knew already: the children's mental health is either already heavily affected by their parent's

disease or is threatened to be affected.

There seems to be nothing wrong with this case either. Before the study, Sam knows that the children's mental health is affected. The proposition is true, Sam believes it and she has sufficient evidence through her own experience. Moreover, Sam's knowledge is also sufficient for her to appropriately treat the children. Even without clinical experience, we can easily see that a parent's depression is very likely to negatively affect the children's mental health. From a *scientific* point of view, however, running a study is necessary. Imagine, again, the psychology department is planning to edit a new textbook, and they ask Sam to write a contribution on the affects of a major depression on the patient's children. It would be unacceptable for Sam to publish her knowledge acquired while talking to her patients in a textbook. In order to meet scientific standards of evidence, she clearly needs to run the study.

As I have presented them, the cases seem to be ones in which a subject knows a proposition that P but is not in a good enough epistemic position to assert that P in a scientific context. The reason seems to be that the subject has only non-scientific evidence, i.e., evidence not generated through recognized scientific methods. The subject does not have the right kind of evidence.

This description of the cases is open to several possible objections. First, some might deny that the non-scientific evidence the subject has for P is enough for knowledge that P . Second, even if we accept that *Sally the Teacher* and *Sam the Psychologist* are cases in which a subject knows that P but is not in a good enough epistemic position to assert that P in a scientific context, there is a question of what explains why this is so. For instance, our cases constitute potential counterexamples to the sufficiency direction of the knowledge norm of assertion. The knowledge norm of assertion states that one is properly epistemically positioned to assert P if one knows that P . Further, one might think that it also undermines the sufficiency direction of the knowledge norm of practical reasoning. The knowledge norm of practical reasoning states that one is properly epistemically positioned to use P in practical reasoning if one knows

that P .⁶ Third, contextualists or subject sensitive invariantists may suggest that the subject has enough evidence for knowledge in ordinary contexts but not in scientific contexts where the standards are higher. Thus, they may offer an interpretation of the cases on which in the scientific context when the subject is not epistemically well enough positioned to assert the relevant proposition, or act on it, she doesn't know it. I will deal with each of these objections in turn.

5.3 Objections and Replies

The following Objections 1 to 5 provide some plausible alternative explanations of why it seems inappropriate for Sally and Sam not to run a study before publishing a paper, which I will all reject.

Objection 1

A first response might be to doubt that the cases are described correctly. One may think that Sally does not know that bullies choose victims that are already rejected by their classmates before having run the study, and that Sam does not know that a major depression of a parent negatively affects the children's mental health before having run the study. According to this objection, Sally may still be properly epistemically positioned to take precautions in the classroom, and Sam may still be properly epistemically positioned to treat the patient's children, despite having no knowledge. However, one may say, knowledge is necessary for Sally and Sam to assert their beliefs in a *scientific context*, i.e., to publish a paper.

In reply and to defend the claim that Sally and Sam have knowledge, the following can be said. First, given their evidence, it is *prima facie* plausible to think that Sally and Sam have knowledge. Take the online articles quoted above again: it is at least understandable why someone unfamiliar with scientific

⁶The knowledge norm of assertion or of practical reasoning is often used or defended as a bi-conditional: a subject knows that P if and only if she is properly epistemically positioned to assert P and rely on P in her practical reasoning.

standards might make fun of the studies. It seems that the researchers already knew what they then confirmed with scientific methods. Similarly, any subject in a similar situation as Sally and Sam seems to have knowledge of the relevant proposition before running a scientific study.

Second, we can easily adapt the cases such that the subject is in an even better epistemic situation which makes it even more plausible that she has knowledge. Here is a slightly changed version of *Sally the Teacher*.

Sally the Experienced Teacher

Sally is a teacher. She has been teaching for a couple of years and has been observing a lot of bullying in her classroom. Sally comes to believe that bullies choose victims that are already rejected by their classmates, which gives them a rise in status on the one hand and no loss in affection by other classmates on the other hand. She speaks to many teachers in her school that have observed similar situations and confirm Sally's observations about bullying. *Sally teaches for 10 more years in different schools. Over the years, she makes the same observations again and again. Because she is very interested in the topic, she talks to many teachers and students, and they all confirm her observations.* Given her evidence, Sally knows that bullies choose victims that are already rejected by their classmates, and she acts appropriately when she takes precautions to protect potential victims from potential bullies. Sally goes back to university and carries out a study that confirms exactly what she already knew: bullies want to realize both status and affection and therefore choose victims that are already disliked by their classmates.

Given that Sally is epistemically extremely well positioned, it is very plausible that she has knowledge. Simply judging from the number of samples, she is in a much better epistemic position before running the study than she

would be from only running the study, without having any previous experience. However, while she can publish her evidence gained through the study in a psychology textbook, she cannot appropriately publish her evidence gained in the classroom and through talking to teachers and students.

If one does not want to grant Sally knowledge in *Sally the Experienced Teacher*, it seems one would have to argue on more general theoretical grounds, such as that we cannot know any general empirical proposition before having tested it scientifically. This, however, would be very implausible. For instance, I know that tulips need water to grow, that most dogs enjoy running and playing with a stick, that people with friends are usually happier than people without friends, etc. I know these general empirical propositions without having tested them scientifically myself and without knowing of a scientific study in which they were tested, and it is likely that some of them have never been scientifically tested at all.

Surely, Sally and Sam cannot rule out that they are deceived by all students, co-teachers and patients, respectively, but in many cases of knowledge we are not able to rule out being deceived, even after having tested our beliefs scientifically. We do not need to assume that in order for a subject to have knowledge, she can rule out these sceptical possibilities.⁷

Objection 2

Someone might question my diagnosis that the reason why Sam and Sally cannot appropriately publish their knowledge is that they do not have the right kind of evidence. Instead, they may suggest that the problem is not a certain quality of evidence they lack but the quantity. To take *Sally the Teacher* as an example: even though Sally can act on her knowledge when she protects disliked students, she is not in a good enough epistemic position to appropriately assert her knowledge in a scientific context because she simply needs *more* evidence.

⁷See, e.g., Williamson [137].

Take *Sally the Experienced Teacher* again and change it such that we grant Sally knowledge from the beginning on: given her evidence, Sally knows that bullies choose victims that are already rejected by their classmates, and she acts appropriately when she takes precautions to protect potential victims from potential bullies. Sally then teaches for 10 more years, makes the same observations again and again, and talks to many teachers and students who all confirm her observations. Even though Sally is epistemically extremely well positioned, far better than is required for knowledge, her evidence is clearly still not sufficient to write a paper on bullying. The reason, I take it, is that it is not the right kind of evidence to be published in the psychology textbook.

Objection 3

My suggestion that the cases of Sally and Sam are ones in which a subject knows that P but is not in a good enough epistemic position to assert it may seem to violate the knowledge norm of assertion or practical reasoning. According to the sufficiency direction of the knowledge norm of assertion, one is properly epistemically positioned to assert P if one knows that P . According to the sufficiency direction of the knowledge norm of practical reasoning, one is properly epistemically positioned to use P in practical reasoning if one knows that P . Denying the knowledge norm of assertion or practical reasoning would be a controversial move since it is widely accepted (e.g., by Williamson [137], Hawthorne [51], Stanley [116]).

However, evidence against the knowledge norm of practical reasoning has recently been given by Jessica Brown [9] and Jennifer Lackey [69], [70]. Brown argues that the knowledge norm cannot serve to settle debates about knowledge, and the cases she gives suggest that knowing that P is not sufficient to appropriately rely on P in practical reasoning. Lackey [70] argues that assertions grounded in a very high degree of justification and even above the threshold for knowledge can fail to be epistemically appropriate and hence provide putative counterexamples to the sufficiency direction of the knowledge

norm. To show this, Lackey gives several cases in which a subject has what she calls ‘isolated secondhand knowledge’: the subject knows that *P* solely on the basis of testimony that *P* and knows nothing or very little relevant about the matter other than that *P*. In these cases, the subject cannot appropriately assert her knowledge that *P*. Lackey claims that the reason is that *P* is not the right kind of evidence. This is similar to what I claimed for my cases: before the study, Sally and Sam lack evidence generated through recognized scientific methods, which is the kind of evidence they need in order to publish their knowledge. While nothing depends on it, I take it that publishing a science paper is an elaborate form of asserting rather than acting.

Let us have a closer look at one of Lackey’s cases:

DOCTOR: Matilda is an oncologist at a teaching hospital who has been diagnosing and treating various kinds of cancers for the past fifteen years. One of her patients, Derek, was recently referred to her office because he has been experiencing intense abdominal pain for a couple of weeks. After requesting an ultrasound and MRI, the results of the tests arrived on Matilda’s day off; consequently, all of the relevant data were reviewed by Nancy, a competent medical student in oncology training at her hospital. Being able to confer for only a very brief period of time prior to Derek’s appointment today, Nancy communicated to Matilda simply that her diagnosis is pancreatic cancer, without offering any of the details of the test results or the reasons underlying her conclusion. Shortly thereafter, Matilda had her appointment with Derek, where she truly asserts to him purely on the basis of Nancy’s reliable testimony, ‘I’m very sorry to tell you this, but you have pancreatic cancer.’ [70, pp. 3-4]

According to Lackey, Matilda knows that Derek has pancreatic cancer, but she is not epistemically well enough positioned to assert that Derek has pancreatic cancer. Isolated secondhand knowledge is simply not the right kind of

evidence to be appropriately used in this context, and hence the case violates the sufficiency direction of the knowledge norm of assertion. My suggestion that the cases of Sally and Sam are ones in which a subject knows that P but is not in a good enough epistemic position to assert P may suggest that they violate the knowledge norm of assertion in exactly the same way.

Lackey discusses an objection to her view that DOCTOR and similar cases violate the sufficiency direction of the knowledge norm according to which Matilda cannot appropriately assert that Derek has pancreatic cancer because this would violate an *institutional* norm. While Matilda cannot be criticized qua asserter, she can be criticized qua doctor: as a doctor, she is not epistemically well enough positioned to assert that Derek has pancreatic cancer. According to this objection to Lackey's case, the medical profession requires that doctors offer diagnoses which are not grounded entirely in isolated secondhand knowledge, in which case Matilda is criticizable only *for institutional reasons*. If this is correct, DOCTOR is consistent with the knowledge norm of assertion. Lackey argues that this objection is not successful for two reasons. First, if the medical profession allowed diagnoses grounded in isolated secondhand knowledge, the institution would no longer serve their epistemic purpose and patients would no longer trust it, so the reasons why Matilda cannot assert her knowledge are epistemic. Second, this explanation does not work for cases which are in relevant aspects similar to DOCTOR but where no institution is involved.⁸

Whether or not Lackey has established that DOCTOR and similar cases violate the sufficiency direction of the knowledge norm of assertion does not matter for the current purposes. I do not think that my cases violate it, but Lackey's cases can nevertheless be used to serve my purpose whether or not they are treated as counterexamples to the knowledge norm of assertion. The reason is that the relevant similarity between Lackey's cases and our cases is that the subject does not have *the right kind of evidence*. Given Lackey's cases

⁸For such cases, see Lackey [70, pp. 22-26].

are very similar to our cases, some of her ways of defending them will work for our cases as well (see Objections 4 and 5).

The weaker claim, that in DOCTOR and similar cases, a proposition P cannot be asserted for *institutional* reasons supports my view. It seems (and the real cases described above suggest) that in a scientific context, knowledge can only be appropriately asserted (e.g., published) when it meets standards such as statistical relevance, objectivity, and reproducibility, and when the methods by which it is gained are made transparent for reasons dictated by the scientific community. From an epistemic point of view, however, it might be perfectly appropriate to assert P in a scientific context. *Sally the Teacher* and *Sam the Psychologist* therefore seem to be consistent with the knowledge norm of assertion.

Objection 4

One might want to object that the reasons why Sally and Sam cannot appropriately publish their knowledge are not institutional, but for instance merely pragmatic. Sally and Sam, one might want to say, could not possibly write down all their experience in an appropriate way that would convince others and this is why they have to run a study. While this might be true as well, the sciences are clearly an institutionalized group enterprise. Scientists work on projects together and use evidence produced by other scientists in their research. Since they have to be able to rely on each other's evidence, the methods by which the evidence is gained have to be made transparent. In the examples given above, Veenstra et al. give a detailed description of how they reached their conclusion about bullying, and Hedman Ahlström makes the two recognized methods she uses transparent (see section 5.2). Surely, some aspects of it may be pragmatic. For instance, scientists would not be able to reproduce evidence if the methods were not transparent. Other aspects may be epistemic. For instance, scientists want to minimize the chances of making mistakes, and this is why they only use methods that have been proven to be reliable. More-

over, it is important to note that one of the sciences' main aims is the mere acquisition of knowledge. Given that knowledge is one of the main aims, we should expect there to be special constraints on how to produce knowledge.

Since the relevant evidence has to be produced by scientifically recognized methods, I will speak of *methodological* standards of evidence.

Objection 5

A further possible response comes from a contextualist or subject sensitive invariantist view. According to contextualism about knowledge, the truth conditions of knowledge ascriptions depend on the context in which they are uttered, where 'context' refers to features of the knowledge attributor or her situation. According to subject sensitive invariantism, the strength of epistemic position required for knowledge varies depending on the stakes of the subject of knowledge attribution. According to these views, a knowledge ascription could change from being true to being false as a result of an increase in the stakes. Second and more importantly, contextualism holds that the truth conditions of knowledge ascription can be altered not merely by the stakes, but by other factors including salience of error and perhaps institutional standards.

Both could try to explain why it would be inappropriate in *Sally the Teacher* and *Sam the Psychologist* not to conduct the studies by claiming that Sally and Sam have knowledge in the classroom or in the psychological practice, but that they do not have knowledge in the high-stakes context the sciences provide. In reply, I will first argue that the stakes are not necessarily high in the science context (both for the subject and a possible knowledge ascriber), and I will then use one of Lackey's replies to the contextualist or subject sensitive invariantist objection to her cases.

It seems that *Sally the Teacher* is not a case in which the stakes shift from low to high. Two different aspects could raise the stakes in a scientific context. First, the stakes could raise because Sally's knowledge would be available to many people who would act on this knowledge, for which Sally would be

responsible. In the classroom context, the stakes are clearly low. All Sally does is protect potential victims who are anyway disliked by their classmates. Now, it seems that protecting disliked children is good no matter how many people do it—it does not get bad if more people do it. The stakes are therefore not higher, neither for Sally personally nor for Sally as part of the scientific community.

Second, the stakes could raise because Sally would make her knowledge available to the scientific community. One may think that the practical stakes are high because the researcher would risk to ruin her reputation in the scientific community if she published a false claim. However, it happens frequently that scientists publish papers with false claims. In fact, some of the most respected researchers have defended theories that were later proven wrong. To stay in psychology, most of the theories developmental psychologist Jean Piaget has defended have been proven wrong. Nevertheless, Piaget remains one of the most admired psychologists in history, and even though many of his claims were false, these very claims have inspired the discipline and led to extensive studies and improved knowledge of the subject matter. It seems that publishing false claims is not a problem in itself, and the reason is that scientists make the methods they use to gain their evidence transparent in their publications. It would be inappropriate to draw conclusions from the evidence which clearly cannot be drawn, or to publish knowledge based on just observing children in the classroom and talking to fellow teachers. Hence, if the stakes are high for the scientist, they seem to be high because of the way scientific evidence is gained or used and not because the claims made might be false.

Even if what I have just argued is not correct and scientific contexts do, for some reason, necessarily provide high stakes contexts, salience of error, or perhaps institutional standards, it seems that this would not make a contextualist or subject sensitive invariantist explanation true. Here one of Lackey's replies to the contextualist and subject sensitive variantist with respect to DOCTOR and similar cases is helpful. Lackey argues that her cases do not involve high

stakes, but even if they could be construed as involving high stakes, the contextualist's and the subject sensitive variantist's response would be to require simply more evidence for knowledge in these cases. However, Lackey argues that a certain kind of evidence is required.⁹ The same would be true in our cases. I have argued in reply to Objection 2 that simply more evidence is not sufficient; we need a special kind of evidence, namely evidence approved by the scientific community.

I have argued that there are methodological constraints on evidence in the sciences: not any kind of knowledge can be used as evidence in a scientific context. Since I am ultimately interested in evidence in a philosophical context, I will not discuss what exactly these constraints are and instead move on to discuss evidence in philosophy. The main reason why I nevertheless chose examples from the sciences is that I am interested in constraints on evidence *in academic philosophy* rather than in ordinary life situations. However, when we assert or act on knowledge gained in philosophy, we are typically in a philosophical context, which makes it difficult to isolate methodological constraints from epistemic aspects which would apply in ordinary life situations. For instance, it is not clear how one's knowledge that knowledge is not justified true belief or one's knowledge that intuitions cannot be reduced to beliefs would influence how one acts in ordinary life situations. In ordinary life situations, we might be interested in whether some particular cases are cases of knowledge and what intuitions people have, but it is hard to see how a particular definition of either would make a difference as to how we act.¹⁰ We are, however, often not in a scientific context when we assert or act on knowledge which can be tested in the sciences (such as Sally's or Sam's knowledge), which made it easier to isolate methodological constraints from epistemic aspects.

In the next two sections, sections 5.4 and 5.5, I argue that there are method-

⁹Lackey [70, p. 27].

¹⁰Surely, we may act on some knowledge gained through philosophical inquiry. For instance, one might come to know that God exists through philosophical research, which might influence how one acts in ordinary life significantly.

ological constraints on evidence in philosophy and I discuss what they could be.

5.4 Evidence in Philosophy

Even though the methods in the sciences and in philosophy differ, it seems that there are analogous constraints on evidence. As mentioned above, one of the reasons why our evidence in the sciences has to meet certain standards is that the sciences are an institutionalized group enterprise. The same holds for academic philosophy as well: academic philosophy is an institutionalized group enterprise. Just like scientists, philosophers highly rely on each others work, and they often work on projects together. Or they work independently on similar projects, and before publishing a paper, they take the evidence that has already been produced on the relevant topics into consideration. Philosophers therefore aim to make their evidence available to each other. They also often make the methods by which they gain their evidence available. For instance, philosophers do not simply publish the knowledge they gain from a thought experiment, they also present and discuss the thought experiment and explain why they take certain facts to be their evidence.

Moreover, I said that one of the sciences' main aim is the mere acquisition of knowledge. Similarly, one of the main aims of philosophy certainly is the mere acquisition of knowledge. It even seems to be *the* aim of philosophy. While evidence in ordinary life situations often serves mainly practical purposes such as to act on it in ordinary situations, knowledge we use as evidence in academic philosophy is rarely used to act on in ordinary life contexts (as explained above).

We assumed that in thought experiments, we usually gain knowledge of the content of the intuition and knowledge of the fact that we have an intuition. I called this the *Inclusive View* of evidence from thought experiments which Williamson endorses (see section 5.1). Williamson claims that we ought to appeal to the content of the intuition as evidence from thought experiments.

However, I argued in Chapter 4 that appealing to the fact that we have an intuition as initial evidence from thought experiments is a practice we commonly pursue. A constraint on evidence in philosophy could explain why we ought to do so.

In what follows, I discuss what such a constraint could be. I first consider three possible ways Williamson suggests to explain why some misguided philosophers appeal to the fact that they have an intuition and thereby psychologize their evidence: an operational standard of evidence, *Evidence Neutrality*, and *Evidence Neutrality* as a conversational norm. I reject these explanations (this section) and then give my own explanation according to which appealing to the fact that we have an intuition is a practice we ought to pursue (section 5.5).

Rejecting an Operational Standard

In ‘Philosophical ‘Intuitions’ and Scepticism About Judgement’ [138], Williamson claims that one reason why some philosophers psychologize their evidence may be that they think one must always be in a position to know what rationality requires of one. Since it is rational to form one’s beliefs on the basis of one’s evidence, these philosophers will also think that one must always be in a position to know what one’s evidence is. Hence, they endorse an operational standard of evidence. An operational standard of evidence implies that a proposition P constitutes evidence only if we know or are in a position to know that P is part of our evidence.¹¹ According to Williamson, the endorsement of such an operational standard could explain why some philosophers appeal to the fact that they have an intuition rather than to the content of the intuition as evidence from thought experiments:

On any standard of evidence on which our evidence includes the

¹¹To be in a position to know that p entails the following, according to Williamson: ‘[...] it is neither necessary to know p nor sufficient to be physically and psychologically capable of knowing p . No obstacle must block one’s path to knowing p . If one is in a position to know p , and one has done what one is in a position to do to decide whether p is true, then one does know p .’ [137, p. 95]

fact that the subject in a Gettier case lacks knowledge, we are not always in the relevant sense in a position to know what our evidence is. Contrapositively, we are always in a position to know what our evidence is only if our evidence does not include the fact that the subject in a Gettier case lacks knowledge. Thus, for both perception and philosophical judgement, the demand for an operational standard of evidence drives even many non-sceptics to adopt the psychologized standard of evidence to which the sceptics appeal. By that standard, one's evidence is only the fact that it perceptually or intellectually appears to one that P, not the fact that P itself. [138, p. 119]

However, in *Knowledge and its Limits* [137], Williamson famously argues that no operational standard of evidence can be established: there is no general domain of evidence to which we have privileged access (the anti-luminosity thesis).¹² Hence, Williamson claims in the above-mentioned paper, mere knowledge is sufficient to qualify as evidence, in philosophy and in the sciences. He says:

In philosophy as in empirical science, our evidence does not consist of facts with respect to which we are infallible. *Ordinary knowledge is enough.* We have no general guarantee against the possibility that we did not know something that we thought we knew. [138, p. 151, my emphasis]

While it is certainly true that our evidence does not consist in facts with respect to which we are infallible, it does not follow that any kind of evidence is good enough. I argued that there are constraints on evidence to be used in the sciences (sections 5.2 and 5.3), and that evidence in philosophy is likely to be constrained as well (this section). We might, therefore, consider an operational standard as a constraint on evidence *in philosophy only*. Just like a proposition

¹²For Williamson's anti-luminosity thesis, see [137, pp. 93-113; 147-183].

cannot be used as evidence in the sciences if it does not meet the standards of scientific evidence, a proposition might not be usable as evidence in philosophy if it does not meet an operational standard.

This could explain why we do not appeal to the content of our intuition but rather to the fact that we have an intuition as evidence from thought experiments. Even though intuitions do not meet the general operational standard, Williamson grants that we are often in a position to know what intuitions we have:

[It is not the case] that we are not often or typically in a position to know what intuitions we have. The point is just that not even facts about intuition meet the fully operational standard of evidence that was used to exclude other facts. [138, p. 121]

We can say the following:

Operational Evidence. Some of our evidence E is such that we know (or are in a position to know) that we have it whereas other evidence E^* is such that we do not know (or are not in a position to know) that we have it.

Operational Evidence could be read in two different ways. First, it could mean that some kinds of evidence E are *typically* such that we know (or are in a position to know) that we have it, whereas other kinds of evidence E^* are *typically* not such that we know (or are in a position to know) that we have it. For instance, it could mean that the fact that we have an intuition that P is typically such that we know (or are in a position to know) that we know it, whereas P , the content of the intuition, is typically not such that we know (or are in a position to know) that we know it. E , of course, is not necessarily psychological evidence.

Second, Operational Evidence could be read such that from the total evidence a subject has which bears a certain claim, some evidence E is in fact

such that we know (or are in a position to know) that we have it whereas some other evidence E^* is in fact not such that we know (or are in a position to know) that we have it, whatever kind of evidence E^* and E are. For instance, it could mean that in one case, the fact that we have an intuition that P is such that we know (or are in a position to know) that we know it, and in another case, the content of the intuition P is such that we know (or are in a position to know) that we know it. Again, E is not necessarily psychological evidence.

If Operational Evidence is true, we can hold an Operational Norm as a constraint on evidence in philosophy:

Operational Norm*. A proposition P constitutes evidence in philosophy only if it is the kind of evidence for which we usually know (or are in a position to know) that it is part of our evidence.

Operational Norm**. A proposition P constitutes evidence in philosophy only if we in fact know (or are in a position to know) that P is part of our evidence.

Either of these norms could at least explain why we sometimes appeal to the fact that we have an intuition as evidence from a thought experiment. Since the differences will not matter in what follows, I will refer to both versions as the Operational Norm. As a methodological norm for evidence in philosophy, the Operational Norm does not tell us anything about the nature of evidence in general, or about the evidence we have against a certain claim. Instead, it tells us something about which evidence can play a role in philosophy. We thereby distinguish between the epistemic situation a subject is in and methodological constraints on the evidence she can use in her research.

The Operational Norm therefore is compatible with Williamson's general picture of knowledge and evidence. First, we do not have to deny the anti-luminosity thesis, because we only need to presuppose that *some* of our evidence is such that we know we have it, which Williamson grants. Second and closely related, a philosopher defending an operational norm for evidence in

philosophy can stick with Williamson's denial of a general operational standard of evidence.¹³ Third, even though the subject might not be able to use all her evidence which bears on a certain claim in a philosophical context, we can endorse the view that the total evidence a subject has in the case where the content of the intuition is true is different from the total evidence the subject has in the case where the content of the intuition is false.¹⁴ In the case where the content is true and the subject has knowledge, the subject's evidence consists in both the content of the intuition and the fact that the subject has the intuition. In the case where the content is false, it consists only of the fact that the subject has an intuition.

However, mere consistency with Williamson's general epistemic picture is not sufficient to make the Operational Norm plausible within this picture. It seems that there is no motivation for a Williamsonian externalist to endorse an operational standard of evidence in philosophy. Williamson often emphasizes the similarity of philosophy to ordinary thinking.¹⁵ It seems then that there is no motivation for an operational standard of evidence in philosophical thinking, if this thinking is not supposed to be much different from ordinary thinking.¹⁶

I will suggest a weaker methodological constraint on evidence in philosophy in section 5.5. However, I first address two further suggestions Williamson makes in order to explain why some contemporary philosophers misguidedly appeal to the fact that they have an intuition as evidence from thought experiments.

¹³Williamson [138, pp. 116-125].

¹⁴See Williamson [140, chapter 7].

¹⁵For instance with respect to thought experiments: 'The main overall aim is to subsume the epistemology of thought experiments under the epistemology of counterfactual conditionals and metaphysical modality [...] and thereby to reveal it as an application of quite ordinary ways of thinking, not as something peculiarly philosophical.' Williamson [140, p. 180].

¹⁶Sosa [113, pp. 135-153] holds a view similar to the one suggested. He thinks that there are at least two different kinds of knowledge, animal knowledge and reflective knowledge. While animal knowledge does not require that the subject has an epistemic perspective on her believe, reflective knowledge requires that the subject endorses the reliability of the source of her knowledge. In philosophy, we are obviously interested in gaining reflective knowledge.

Rejecting Evidence Neutrality

In *The Philosophy of Philosophy* [140], Williamson claims that philosophers appeal to the fact that they have an intuition because they believe in *Evidence Neutrality*. *Evidence Neutrality* is the thesis that

[...] whether a proposition constitutes evidence is in principle uncontentiously decidable, in the sense that a community of inquirers can always in principle achieve common knowledge as to whether any given proposition constitutes evidence for the inquiry. [140, p. 210]

It is not entirely clear what exactly *Evidence Neutrality* entails. First, it is not clear whether it is meant as an epistemic norm or as a methodological norm, i.e., whether it is meant to concern all our evidence or merely the evidence to be used in philosophical research. Given that the passage mentions a community of inquirers, it at least seems to be a norm which holds in philosophy but not necessarily in other contexts. Second, it is not clear whether *Evidence Neutrality* is meant to say that the inquirers ought to be able to gain common knowledge as to whether a proposition is evidence (and hence, as to whether it is knowledge), or whether the inquirers ought to be able to gain common knowledge as to whether a proposition is true. The first would be an even stricter constraint than the Operational Norm, adding the requirement that the community be in a position to agree on the operational evidence. Third, it is not clear what it means that a community of inquirers can always ‘in principle’ achieve common knowledge. If we take ‘in principle’ to mean anything substantial and not to be trivial (such as that in principle, we could agree on every single proposition), it should be understood such that the community of inquirers ought to either agree or be in a position to agree on what the evidence is. This corresponds to how Brian Weatherson [132] seems to understand *Evidence Neutrality*:

Neutrality Norm. A proposition P constitutes evidence in philosophy only if the philosophical community agrees that P is true.¹⁷

Weatherson defends the Neutrality Norm as a methodological norm. He claims that not only in philosophy, but also in the sciences ought we to appeal only to evidence which is agreed upon by all parties of a debate. A subject who follows the Neutrality Norm, Weatherson argues, will acquire more knowledge and contribute to more knowledge of the community than a subject who does not.

The Neutrality Norm, however, seems to be a rather unrealistic methodological constraint. Even within a non-sceptical philosophical community, philosophers radically disagree on many topics, and using only evidence that everyone agrees on would reduce our evidence drastically. Moreover, philosophical dispute often takes on a conditional form. We do not necessarily endorse the premisses of our opponents, but we engage in conditional reasoning on the basis of our opponents' premisses. Having to agree on every single step would not get anyone anywhere, and progress is often made through working out theses on the basis of claims we do not endorse. Hence, the Neutrality Norm does not correspond to our practice and hence cannot explain our practice, so even if for some reason the Neutrality Norm would be ideal and lead to more common knowledge, we would constantly violate it. We are, however, looking for a norm which can explain our practice.

Rejecting Conversational Norms

One might want to argue that other than in the sciences, we can explain our practice in philosophy on a conversational level. While I take methodological norms to concern specifically the sciences and philosophy, conversational norms are more general and apply in any context in which we communicate.

¹⁷Weatherson [132] points out that this claim is ambiguous. It could either mean that a community can in principle agree that P is true, or it could mean that a community can in principle agree that P bears on a certain issue. Like Weatherson, I take it to mean the first.

Williamson considers *Evidence Neutrality* as a *conversational* norm and suggests the following:

In debate, one cannot hope to persuade opponents by appealing to evidence that they do not accept. Predictably, they will accuse one of begging the question. A fact can function as evidence in the debate only if both sides are willing to accept it. If one party asserts that P while the other party denies that P, they cannot use the fact that P as shared evidence, but they can use the fact that the first party asserts that P as shared evidence, because they presumably agree on that. [138, p. 121]

Even though Williamson thinks that in a conversation, ‘it is dialectically pointless, rude’ to offer as evidence propositions that the opponent rejects, he thinks that we ought not retreat to evidence the sceptic accepts, because a ‘sufficiently ruthless sceptic can challenge everything that we offer as evidence, by always demanding a proof’¹⁸. For this reason, Williamson rejects *Evidence Neutrality* as a conversational norm as well.

To begin with, *Evidence Neutrality* as a conversational norm confronts the same problems as the methodological Neutrality Norm: it does not correspond with our practice, because we simply do not only use evidence everyone agrees on. Furthermore, such an explanation is not a charitable reading as to what philosophers believe, because it presupposes that they have a pretty general false belief about what they do. Not only would their belief be false, it would also be a rather sophisticated meta-philosophical assumption, namely that a community of philosophers can always in principle agree about what constitutes evidence.

If we want to explain why we appeal to the fact that we have an intuition on a conversational level, we can do so on the basis of Gricean maxims, which is more charitable than Williamson’s explanation. In order to establish that our

¹⁸Williamson [138, p. 124, p. 152].

false belief in *Evidence Neutrality* is the reason why we appeal to the fact that we have an intuition, Williamson assumes that facts about psychological states are ‘easier to accept’ than facts about the world.¹⁹ His explanation presupposes at least that we *believe* that they are better accessible than the facts about the world. This is a plausible assumption for the following reason. We rarely give evidence for the claim that we have a certain intuition, which indicates that we are quite confident with respect to whether we have an intuition or not. Moreover, we usually do not question someone else’s claim that they have an intuition, we rather just accept it as true. (I will argue in section 5.5 that it is true that facts about our intuitions are usually easier to access than the contents of intuitions.)

If it is true that we believe that facts about our intuitions are easier to access in the relevant cases, we can analyse the appeal to intuitions from thought experiments by applying some Gricean maxims of conversation. The idea is that there are two pragmatic rules in conflict with each other, the *Maxim of Quantity* and the *Maxim of Quality*.

1*. *Maxim of Quantity*: Make your contribution as informative as is required (for the current purposes of exchange)

2*. *Maxim of Quality*: Do not say that for which you lack adequate evidence

Here is an example Grice gives in which the *Maxim of Quantity* is violated, and its violation is to be explained by the supposition of a clash with the *Maxim of Quality*.

A is planning with B an itinerary for a holiday in France. Both know that A wants to see his friend C, if to do so would not involve too great a prolongation of his journey [...]

A: *Where does C live?*

B: *Somewhere in the South of France*

¹⁹Williamson [140, p. 210].

[50, pp. 32-33]

According to Grice, B's answer is less informative than is required to meet A's needs. The violation of the *Maxim of Quantity* is best explained by the supposition that B is aware that being more informative would possibly lead to a violation of the *Maxim of Quality*, since he does not know where exactly in the South of France C lives.

Applied to situations in which we first come up with thought experiments as counterexamples, the *Maxim of Quantity* tells us to appeal directly to the fact about the world as evidence, because this is the best evidence we have. Appealing to the fact about the world as evidence would mean to make the most informative claim possible, because giving the strongest evidence against the theory in question makes the theory most unlikely. However, the *Maxim of Quality* most of the time tells us to appeal to the fact that we have an intuition in such a situation, because we believe that this fact is much easier to access and hence our claim is more likely to be true. Whenever these conditions hold, the conflict between the rules explains why we appeal to our intuitions as evidence. As in the example given by Grice, it is more appropriate to violate the *Maxim of Quantity* than to violate the *Maxim of Quality*.

While I take it that this is a more charitable explanation than Williamson's, there is an objection to any kind of explanation on the conversational level. Philosophical work usually does not take place in ordinary conversation. To engage in philosophical research in its paradigmatic form means to write papers or books for publication within an institutional context. Gricean norms might explain some of our philosophical conversations, but when it comes to explaining our practice of publishing our ideas and arguments in peer reviewed journals, it seems that we have to look at methodological norms and not at norms of conversation, just like in the sciences.

In the next section, I propose a methodological norm which is a natural fit for a Williamsonian epistemic externalist.

5.5 Philosophical Evidence and Accessibility

As mentioned in section 5.4, Williamson assumes that facts about psychological states are ‘easier to accept’ than facts about the world when he argues that our false belief in *Evidence Neutrality* is the reason of why we appeal to the fact that we have an intuition.²⁰ In cases where we disagree with our peers on whether a proposition about the world is evidence, we draw on a fact that is easier to accept by them, namely the fact that someone has an intuition. The fact that someone has an intuition then is supposed to count as shared evidence.

However, to motivate the appeal to the contents of our intuitions, Williamson claims that ‘we frequently have better epistemic access to our immediate physical environment than to our own psychology.’²¹ It is certainly true that we frequently have better epistemic access to facts about the world than to some of our own psychological states. For instance, it is probably true that I have better access to the fact that there is a table in front of me than to why exactly I am in a really good mood today. However, in the case of thought experiments as counterexamples to philosophical theories, it seems quite implausible that we have better access to the content of the intuition than to the fact that we have an intuition. First, thought experiments usually are not about mundane things such as whether there is a table in front of me, but rather about abstract matters of fact such as whether a person has knowledge or only justified true belief. Second, the psychological fact that we have an intuition does not belong to the realm of the unconscious which is difficult to access, as for instance the reason for my good mood might do. Third, even if we can be wrong about our intuitions, the history of the discipline strongly suggests that in philosophy, we are much more likely to be wrong concerning the contents of our intuitions. At least it is obvious that we have to work tremendously hard to get to the facts. We give arguments, examples and counterexamples—and we sometimes realise

²⁰Williamson [140, p. 210].

²¹Williamson [140, p. 5].

only decades later that we were wrong. Although we might occasionally be wrong about ourselves having an intuition or not, it is quite plausible to think that we have better access to the fact that we have an intuition than to the content of our intuition in the case of thought experiments.

Given that in any case, some evidence is better accessible than other evidence, the Williamsonian externalist can accept the following claim about the relative accessibility of facts:

Accessible Evidence. Some facts F are better accessible than other facts F^* .²²

This, again, could mean two things. First, it could mean that some kinds of facts F are *typically* better accessible than other kinds of facts F^* . For instance, it could mean that the fact that we have an intuition that P is typically better accessible than the content of the intuition P . F , of course, is not necessarily internal or psychological in nature, and F^* is not necessarily a fact about the world. Second, it could mean that from the total evidence a subject has for a certain claim, some facts F are in fact better accessible than other facts F^* , whatever kind of facts F and F^* are. For instance, it could mean that in one case, the fact that we have an intuition that P is better accessible than P , and in another case, P is better accessible than the fact that we have the intuition that P . Again, F is not necessarily psychological evidence.

For the purpose of this chapter, nothing depends on how we spell out Accessible Evidence. If there is evidence that is accessible, we can hold the following methodological norm for evidence in philosophy:

Accessibility Norm*. From our total evidence which bears on a certain thesis we ought to appeal to the part which is usually easiest to access.

Accessibility Norm.** From our total evidence which bears on a certain thesis we ought to appeal to the part which is in fact easiest to access.

²²This concerns only facts that are in principle accessible.

Since the differences will not matter in what follows, I will refer to both versions as the Accessibility Norm. Like the Operational Norm, the Accessibility Norm is compatible with Williamson's epistemic views if we take it to be a methodological norm only.

I will now give some reasons why I think the Accessibility Norm explains our practice correctly and has advantages over the Operational Norm or the Neutrality Norm. First, given that philosophy is mainly about the acquisition of knowledge (as I argued in section 5.4), we might want to proceed carefully, and more carefully than in cases where we need to make practical decisions. Since we do not have to act on knowledge gained in philosophy, we can afford to proceed carefully. Rather than risking anything (even in cases where the method is in principle reliable), we better appeal to the fact which is easier to access.

Second, appealing to better accessible evidence will often mean appealing to weaker evidence and force us to acquire more evidence, which then leads to more knowledge. I have shown in Chapter 4 that intuitions from thought experiments as counterexamples to philosophical theories play a role as initial evidence only, which implies that we need to produce more evidence in order to refute a philosophical theory.

Third, from a methodological point of view, it is more reasonable to appeal to evidence which is easier to access in order to establish facts which are harder to access. Sherilynn Roush [102] has recently defended the view that for a proposition E to serve as evidence for a proposition P , it should be easier to access the truth of E than the truth of P .²³ Roush refers to this feature of evidence as 'leverage' and gives the following picture:

[...] knowing that the evidence statement is true is usually a lot easier than knowing that the hypothesis statement is true, and we use the former to help us make progress on the latter where we could not have made progress directly. This is analogous to the

²³Roush [102, chapter 5].

way in which we use a lever—a plank and a well-placed fulcrum—and our ability to lift a small weight in order to move a large weight we could not have lifted directly. The lever would not help us to move weights, of course, if for some reason the plank always had to be as heavy as the larger weight itself. For then we could not set up the connection between the smaller and the larger weight, for the same reason as we could not move the larger weight directly: the plank would be too heavy to put in place. It is similar with evidence. [102, p. 158]

Roush argues that in order to establish the truth or falsehood of a claim we need evidence which bears on the claim, but in order to be useful, it has to be easier to establish the truth of the evidence than the truth of the hypothesis. If we already knew whether the hypothesis was true, we wouldn't need the evidence, just as we would not need a lever if we could already lift the heavy object. However, if we cannot lift the heavy object, then an equally heavy lever would not serve our purposes.

Fourth, it seems plausible that if we appeal to the evidence which is better accessible, we more often know what our evidence is and different parties more often agree on what the evidence is. Hence, while the Accessibility Norm does not entail anything as strong as the Operational Norm and the Neutrality Norm, appealing to evidence that is relatively easy to access often means appealing to evidence that we know we have and that others agree with us on.²⁴ Hence, following the Accessibility Norm is conducive to both the idea that we should know what our evidence is and the idea that we should be able to share our evidence with our peers, even if as methodological constraints, these requirements would be too strong.

²⁴Williamson says that Williamson says that [o]f course, we can often decide whether a proposition constitutes evidence prior to deciding the main issue, otherwise the notion of evidence would be useless', [140, pp. 212-213].

5.6 Conclusion: Appealing to Intuitions and Scepticism

Let me summarize what I argued in Chapters 3, 4 and 5 (the present chapter). In Chapter 3, I argued that attempts to show that intuitions do not play a role as evidence in philosophy fail. In Chapter 4, I argued that using the fact that we have an intuition as evidence from a thought experiment is a common practice. To show that we do so for methodological reasons, I argued in this chapter that there is a methodological constraint on evidence in philosophy. I first showed that there are methodological constraints in the sciences (sections 5.2 and 5.3). Given that academic philosophy is similar to the sciences in relevant aspects, there are likely to be methodological constraints on our evidence in philosophy as well. I considered several possible methodological norms and then suggested that a norm according to which we ought to use evidence that is relatively easy to access can explain why we appeal to the fact that we have an intuition as evidence from thought experiments even if we endorse the *Inclusive View* (sections 5.4 and 5.5).

I mentioned in Chapter 3 that Williamson [140] thinks appealing to psychological facts opens a gap which is not easy to close and causes scepticism.²⁵ I will now show how what I have argued in chapter 4 and in this chapter addresses this problem.

As a first reply to the worry of scepticism caused by the appeal to facts about our psychology, note that *prima facie*, there is nothing special about the gap between these facts and facts about the world: we often appeal to the fact that some theory is simple as evidence for its truth, we often use semantic facts as evidence for metaphysical facts, and we use syntactic facts as evidence for semantic facts. These strategies all open a gap which has to be closed and, according to many philosophers, can in fact be closed.

Here are some examples in more detail. Most scientists as well as philosophers think that, other things being equal, the fact that a theory T is simpler

²⁵Williamson [140, p. 234].

than an alternative theory T^* is a reason to choose T over T^* .²⁶ According to Elliott Sober, philosophers take the simplicity of a theory to be evidence for its truth:

Scientists sometimes choose between rival hypotheses on the basis of their simplicity. Non-scientists do the same thing; this is no surprise, given that the methods used in science often reflect patterns of reasoning that are at work in everyday life. When people choose the simpler of two theories, this ‘choosing’ can mean different things. The simpler theory may be chosen because it is aesthetically more pleasing, because it is easier to understand or remember, or because it is easier to test. However, when philosophers talk about the ‘problem of simplicity’, they are usually thinking about another sort of choosing. The idea is that choosing the simpler theory means regarding it as more plausible than its more complex rival. [109, p. 13]

To take a second example, one of Williamson’s [140] concerns is to make clear that the (main) subject matter of philosophy is neither language nor concepts. Philosophy is as much about the world as any other science. Nevertheless, Williamson thinks that it is sometimes useful to study the meaning of our words and the semantic structure of our sentences when we are interested in metaphysical facts. Since in philosophy, we have to reason ‘in areas where it is very hard to distinguish valid from invalid reasoning’, we must pay attention to the semantic form of the premises, conclusion and intermediate steps, which requires semantic beliefs and sometimes explicit testing of these beliefs. Because semantic theories ‘have the great methodological advantage of being comparatively easy to test in comparatively uncontentious ways’ and hence are easier to test than facts about the world, we sometimes have to engage in

²⁶Simplicity can be understood in many ways. It may stand for the paucity of parameters, for theoretical plausibility or elegance, communicability, metaphysical parsimony and many other things. See Baker [5], Zellner et al. [143].

semantic theorizing.²⁷ The semantic facts can then be used as evidence for metaphysical facts, according to Williamson.

The attempt to provide a semantic theory that coheres with a given metaphysical claim can therefore constitute a searching test of the latter claim, even though semantics and metaphysics have different objects. [140, p. 285]

As a third example, take facts about the structure of ordinary language on the one hand and semantic facts on the other hand. Philosophers sometimes take syntactic facts to be evidence for semantic facts, for instance in the debate about the nature of propositional attitudes. Throughout the history of philosophy, the available views were variations of relational views of propositional attitudes. According to these views, propositional attitudes are relations between a subject and a propositional content (be it an abstract object as according to Frege and Russell or a psychological object in a Fodorian language of thought). The problem with relational views is that they do not meet all widely accepted adequacy conditions for propositional attitudes. Quine and others have come up with non-relational dispositional or functional accounts of propositional attitudes. These views, however, face the difficulty that they do not match the structure of ordinary language attitude attributions. According to ordinary language, we can conclude from the fact that Jones believes that Bill will be late both that there is something that Jones believes and that there is someone who Jones believes will be late. Non-relational views cannot account for this fact, because they cannot account for the fact that we can quantify over the predicate's that-clause and over the subject position of the embedded sentence.²⁸ Robert Matthews summarizes the non-relational view and its main challenge as follows:

What looks to be a relational predicate, they argue, is in fact a monadic predicate in which the constituent terms are semantically

²⁷Williamson [140, p. 284-285].

²⁸See Matthews [83, p. 1-2].

‘fused’, such that we can imagine these terms being linked by invisible hyphens (e.g., believes-that-Bill-will-be-late) [...] if we are going to be non-relationalists about the attitudes [...] then we are going to have to find some way of abandoning the assumption shared by relationalists and non-relationalists alike, that the adicity of propositional attitudes must match the adicity of the predicates by which we standardly attribute them. [83, p. 1-2]

It seems to claim that we should not appeal to facts about our psychological states because there is a gap between these facts and the content of our psychological states is not sufficient to make the case against this practice. There is a similar gap between facts about theoretical virtues, semantic facts, or syntactic facts on the one hand and facts about the subject matter we are actually interested in on the other hand. However, we often appeal to facts about theoretical virtues, or to semantic or syntactic facts as evidence.

Closing the gap between evidence that is easier and evidence that is harder to access might not be that hard. However, we should not expect one general strategy to do the work here, because the kinds of evidence that are easier to access as well as the kinds of subject matter our theories are about differ. It seems that each link needs to be defended or proved unreliable independently (they might run parallel of course).

Whether semantic facts mirror metaphysical facts and whether the structure of language tells us something about its semantics are both ongoing subjects of discussion in philosophy. As an example of the second, I mentioned the debate about propositional attitudes. As to simplicity (both parsimony and elegance), philosophers have been giving both *a priori* and *a posteriori* arguments as to why, other things being equal, the simpler theory is more likely to be true.²⁹ Amongst the *a priori* justifications there are theological (e.g., Smart [107]), and metaphysical (e.g., Lewis [73]) justifications and such that claim that there is intrinsic value in simplicity (e.g., Sober [109]). However,

²⁹For a summary of arguments, see Baker [5].

philosophers have also given naturalistic justifications. The important point to note is that philosophers have been trying to close these gaps, which requires philosophical work.

Let us see how the gap between facts about our psychological states and the contents of these states could be closed. This case should be less controversial than any of the cases mentioned above. The reason is that if our intuitions (e.g., from thought experiments) are reliable, we are both reliable at getting the fact that P right as well as at having an intuition that P , because whenever we get a fact that P right (e.g., in the case of thought experiments) we get it right by having an intuition. In other words, whenever we know that P in these cases, it is also true that we have an intuition that P . Hence, any argument for why we are reliable in judging in a certain area or as part of a certain method is also good as an argument for why we can take the fact that we have an intuition as evidence. Whether we are reliable has to be shown separately for any area or method, it certainly does not generalize to intuitions in general, but I was concerned with intuitions in thought experiments only.

Jessica Brown [10] argues that we can close this gap within an externalist picture of epistemology, which I mentioned in Chapter 3. Brown refers to perceptual evidence and claims that we might be able to use the same strategies in the case of intuitions. Brown compares the gap we are concerned with with a gap between psychological evidence from perceptual experiences and facts about the world. Just as we can close the gap between perceptual experience and facts about the world, we can close the gap between facts about our intuitions and facts about the world by appeal to an externalist approach to justification and/or knowledge:

Suppose that whenever one has such an experience, one forms an appropriately related belief. For instance, that when one has an experience as of a large barking dog in front of one, one forms the belief that there is a large barking dog in front of one. On an externalist approach to justification, such as reliabilism, as long as

the appropriate external relations hold, the beliefs so formed are justified. One might hope to apply this solution to the gap problem in the case of perception to the gap problem facing the psychological view of thought-experiment evidence. Suppose that, in fact, the method of forming beliefs about the nonpsychological subject matter of philosophy on the basis of the relevant psychological propositions is reliable. Combining this supposition with a reliabilist approach to justification has the result that beliefs formed in this manner are justified. [10, p. 513]

I argued in sections 5.4 and 5.5 that there is a methodological constraint on evidence in philosophy which explains why we appeal to the fact that we have an intuition as evidence from thought experiments even if we endorse what I called the *Inclusive View*: the Accessibility Norm. Note that the Accessibility Norm could also explain why we often appeal to facts about theoretical virtues, to semantic or syntactic facts. It seems that it is easier to access facts about the simplicity of a theory than to access the truth of the theory. Similarly, it seems easier to access facts about semantics than metaphysical facts, and it seems easier to access facts about the structure of ordinary language than to access semantic facts. To retreat to pieces of our total evidence that are typically or in fact relatively easy to access is a common strategy in philosophy.

I presupposed that we do not have to endorse the sceptical consequences some experimental philosophers draw from their studies and that philosophers' intuitions are more reliable than laypeople's intuitions (see Introduction). If we appeal to the fact that we have an intuition, however, we have to argue from this psychological fact to a fact about the world. Since we appeal to the fact that we have an intuition for methodological reasons only, it follows that we have to close the gap for methodological reasons only. This, however, does not affect our knowledge, and hence does not constitute a sceptical threat.

While we know that a thought experiment undermines the target theory, we might not have succeeded in refuting the theory. I showed in Chapter 4 that

there are several ways to challenge an intuitive counterexample and to argue that a particular intuition does not lead us to the truth or is not inconsistent with the target theory. I think that for methodological reasons, we have to show that these strategies do not apply for any particular case we are concerned with. It might, therefore, be hard to show that a counterexample is not merely intuitive but instead undermines a certain theory. This does not seem to come as a surprise. The history of philosophy proves that it is very hard to refute a philosophical theory.

Chapter 6

Reflective Equilibrium and Counterexamples

Abstract

Reflective equilibrium has been considered a paradigm method involving intuitions. Some philosophers have recently claimed that it is trivial and can even accommodate the sort of scepticism about the reliability of intuitions advocated by experimental philosophers. In order to see whether it is as trivial as some philosophers think, I discuss whether reflective equilibrium is compatible with different views concerning our practice of appealing to intuitions as evidence from thought experiments. I first argue that reflective equilibrium is not compatible with scepticism about the reliability of intuitions. I then argue that reflective equilibrium is compatible with the view I have defended in Chapters 3 to 5, but that we have to make some more specific methodological claims to account for it.

6.1 Introduction

Reflective equilibrium has been considered a paradigm philosophical method involving intuitions. It has been extensively discussed in normative ethics and political philosophy. The key idea of reflective equilibrium as introduced by

John Rawls [99], [100] for moral and political philosophy is that we test our moral judgments (or intuitions) and moral principles against each other and revise and refine both when they are inconsistent.¹ In the literature, ‘reflective equilibrium’ ambiguously stands for the *method* of reflective equilibrium and for the *state* in which all our beliefs are in equilibrium, i.e., in which we have reached coherence amongst our beliefs. I am concerned with the method of reflective equilibrium (MRE in what follows).

Without specifying what exactly they mean by it, philosophers in all areas of research frequently use the term ‘reflective equilibrium’ when they mention the methods and aims of their inquiry. It has been suggested that ‘reflective equilibrium’ is nothing more than a metaphor for the rational performance of philosophy: for taking into account all relevant information available and for working out the most plausible, coherent, and comprehensive theory of the subject matter under investigation.² Michael DePaul [25] and Peter Singer [104] suggest that it could even be compatible with scepticism about the reliability of intuitions, i.e., with a view according to which we ought not take our intuitions into account at all.

I am interested in the question whether MRE is as trivial as some philosophers think or whether it gives us some helpful methodological advice. If it gives us methodological advice, having followed reflective equilibrium could provide us with some evidence as to the truth of the resultant theory which is relatively easy to access and hence is evidence we ought to appeal to in order to establish this theory (as argued in Chapter 5). Different views concerning the role of intuitions that have been defended in the recent debate on the methodology of philosophy have not yet been reflected in the literature on MRE. The current chapter looks at MRE in particular in light of the distinction between the appeal to the content of an intuition and the appeal to the fact that we

¹Even though the term ‘reflective equilibrium’ was introduced by Rawls, Nelson Goodman was the first to discuss the method behind the name in ‘The New Riddle of Induction’ (originally 1953) in his *Fact, Fiction, and Forecast* [47]. Goodman defends reflective equilibrium as a method of justification in logic and inductive reasoning; see also Resnik [101].

²E.g., DePaul [24], [25], Foley [36], Singer [104], Williamson [140].

have an intuition.

The chapter consists of two main parts. In the first part (sections 6.2 to 6.4), I present MRE as it has been discussed in moral philosophy (section 6.2) and apply it to epistemology, more precisely to Sophie, a fictional philosopher who thinks about the Gettier Cases as counterexamples to the theory of knowledge as justified true belief (the JTB theory of knowledge) (section 6.3). I then specify what philosophers mean when they say that MRE is trivial (section 6.4). In the second part (sections 6.5 and 6.6), I look at different views philosophers hold in the current debate concerning the role of intuitions in philosophy. I argue that the sceptical view according to which intuitions ought not play any role in philosophy is not compatible with MRE (section 6.5). I moreover argue that the view I defended in Chapters 3 to 5, according to which we ought to appeal to the fact that we have an intuition as initial evidence from thought experiments is compatible with MRE, but more specific methodological claims have to be made in order to distinguish it from other views and in order for MRE to serve as a methodological guide (section 6.6). I conclude that MRE is not as trivial as some philosophers think (section 6.7).

6.2 The Method of Reflective Equilibrium

Let us look at what exactly MRE amounts to according to some moral philosophers. Two versions of MRE have been distinguished, *narrow* and *wide* MRE.³ According to narrow MRE, we take (a) a set of considered moral judgments (i.e., judgments made or intuitions had in certain circumstances conducive to the truth of their content) held by a particular person and (b) a set of general moral principles and produce a coherent theory by adjusting either (a) or (b), or both. In wide MRE, we extend the area of considered judgments and principles we take into consideration to reach coherence among our *widest set of beliefs*. The following is the general idea which lies behind wide MRE,

³Daniels finds this distinction implicit in Rawls [99] and explicit in Rawls [100], see Daniels [19, p. 257, footnote 2].

according to Norman Daniels:

The method of wide reflective equilibrium is an attempt to produce coherence in an ordered triple of sets of beliefs held by a particular person, namely, (a) a set of considered moral judgments, (b) a set of moral principles, and (c) a set of relevant background theories. We begin by collecting the person's initial moral judgments and filter them to include only those of which he is relatively confident and which have been made under conditions conducive to avoiding errors of judgment. For example, the person is calm and has adequate information about cases being judged. We then propose alternative sets of moral principles that have varying degrees of 'fit' with the moral judgments. We do *not* simply settle for the best fit of principles with judgments, however, which would give us only a *narrow* equilibrium. Instead, we advance philosophical arguments intended to bring out the relative strengths and weaknesses of the alternative sets of principles (or competing moral conceptions). These arguments can be construed as inferences from some set of relevant background theories [...] [19, p. 258]

Wide MRE obviously shares some features with narrow MRE. As with narrow MRE, we determine a set of considered moral judgments and the best-fitting principles. However, we also propose alternative sets of principles and use some of our background theories to provide arguments for and against our competing sets of principles. We finally choose between the different sets of principles based on arguments from the background theories. In this process, we are prepared to adjust the considered judgments, the principles, and the background theories.⁴

Rawls, Daniels, and other proponents of MRE in moral and political philosophy think that the inclusion of alternative sets of principles and background

⁴See also Daniels [20, p. 25].

theories is essential to a method in these areas.⁵ My focus lies on the question of how MRE applies to areas such as epistemology, metaphysics, or philosophy of language, where the subject matter under investigation does not consist of moral or similar norms. In these areas at least, it seems that background theories play a role when we solve conflicts between our intuitions and beliefs. Narrow MRE will therefore not be discussed any further, and I will use ‘MRE’ for wide MRE in what follows.

I will focus on Daniels’ [19], [20], [21] formulation of MRE. It can be given in several steps, which will facilitate the application of MRE to our practice of debating thought experiments as counterexamples to philosophical theories. I will take ‘intuitions’ as synonymous with ‘initial moral judgments’ and ‘considered moral judgments’, and I will take theories to be the counterparts of moral principles in non-moral philosophy.

Step 1. *The relevant intuitions.* Amongst all our intuitions, we choose a set of intuitions of which we are relatively confident, and which we had under ideal conditions, such as having adequate information about the subject matter under investigation and being in a state of mind that is not conducive to error.

Step 2. *The best-fitting theory.* We determine a theory that fits best with the set of intuitions, i.e., that is directly supported by the set of intuitions.

Step 3. *Alternative theories.* We determine alternative theories that are not directly supported by the set of intuitions, but do concern the subject matter under investigation. In the case of a counterexample to an accepted philosophical theory, our currently accepted theory will be amongst these theories.

Step 4. *The relevant background theories.* We look for empirical and philosophical background theories that deliver arguments for

⁵See, e.g., Rawls [99, p. 49]. Nelson Goodman [47], in contrast, defends narrow MRE for the justification of inductive and deductive forms of reasoning.

or against the competing theories.

Step 5. *Restoring coherence.* We use arguments from our background theories and our intuitions to figure out the best and most coherent theory by either disregarding our intuitions, adjusting our accepted theory, or adjusting our background theories, or all three.

In the next section, section 6.3, I will give an example to see how exactly steps 1 to 5 of MRE could apply to thought experiments as counterexamples to philosophical theories.

6.3 Reflective Equilibrium Applied to Counterexamples

Whenever some philosopher comes up with an intuitive counterexample to an accepted philosophical theory, we confront the task of resolving the resultant inconsistency in one or the other way. MRE could be the method we ought to apply to regain consistency. I will go through the steps of MRE as presented above and apply them to a well-known case of inconsistency in epistemology: the JTB theory of knowledge and the Gettier intuitions. I will refer back to this case when I discuss what philosophers mean when they say that MRE is trivial and whether it is compatible with different views on our practice of debating thought experiments as counterexamples to philosophical theories (sections 6.4 to 6.6).

Step 1. *The relevant intuitions*

Let us imagine an epistemologist Sophie who lives in the year 1963. Sophie has been thinking about knowledge for a long time, and one day she gets to read Edmund Gettier's paper 'Is Justified True Belief Knowledge?'. Here is one of Gettier's cases (for present purposes, Gettier's second case is sufficiently similar that we do not need to describe it as well):

Suppose that while Smith has strong evidence that his friend Jones owns a Ford, he has no idea where his friend Brown is. Smith

randomly selects three place-names and constructs the following three propositions: either Jones owns a Ford or Brown is in Boston; either Jones owns a Ford or Brown is in Barcelona; either Jones owns a Ford or Brown is in Brest-Litovsk. Even though Smith has no idea where Brown is, he is justified in believing each of these three propositions, because he has correctly inferred them from a proposition for which he has strong evidence, namely that Jones owns a Ford. However, Jones does in fact not own a Ford, but is driving a rented car. Unknown to Smith, Brown happens to be in Barcelona. It seems that Smith does not know that either Jones owns a Ford or Brown is in Barcelona, even though it is true, Smith believes that it is true, and Smith is justified in believing that it is true.⁶

Sophie agrees with the description of the case. She has the intuition that Smith does not know that either Jones owns a Ford or Brown is in Barcelona. However, the theory of knowledge that Sophie holds, the JTB theory, predicts that Smith knows that either Jones owns a Ford or Brown is in Barcelona. Hence, the content of Sophie's intuition is clearly inconsistent with what follows from the JTB theory.

Suppose that Sophie notices the inconsistency and strives to resolve it. Suppose furthermore that Sophie's intuition meets the relevant criteria: Sophie is relatively confident of the intuition, and she had it under ideal conditions. Her intuition therefore qualifies for MRE (in Daniels' terms, the intuition is a 'considered judgment' as opposed to an 'initial judgment').

Step 2. *The best-fitting theory* and Step 3. *Alternative theories*

Sophie finds the JTB theory supported by a set of her intuitions about cases where justified true belief seemingly is knowledge. For instance, Sophie has the following intuition: if she hears her cat meowing in the kitchen she is justified

⁶Gettier [41, p. 14-15].

in believing that her cat is in the kitchen, and if it is also true that her cat is in the kitchen, then she knows that her cat is in the kitchen. Her intuitions in the Gettier Cases are not part of this set of intuitions, and Sophie now has to determine the theory that fits her Gettier intuitions best as well as alternative theories that compete with the best-fitting theory.

Sophie reads Michael Clark's response to Gettier, 'Knowledge and Grounds: A comment on Mr. Gettier's Paper'. She spends a long time thinking about the cases and about Clark's no-false-lemma reply to Gettier. In both Gettier Cases, Smith's reasoning is based on a false premise. In the case quoted above, Smith gains his justified true belief by reasoning from the justified false belief that Jones owns a Ford. The case Clark gives in reply has it that knowledge is justified true belief where the justification is not based on a false assumption.

Sophie also comes up with her own theories, of which the first is a defeasibility analysis of knowledge. The defeasibility account has it that knowledge is justified true belief absent a defeater.⁷ The second theory is a causal theory, on which the belief that P is knowledge only if it is appropriately causally connected to the fact that P .⁸ The third theory Sophie comes up with involves a simple reliability condition. According to such an account, S knows that P if and only if S 's belief that P is true and justified, where S 's belief that P is justified if and only if the belief that P was produced by a reliable cognitive process.⁹

Sophie decides that Clark's adjusted JTB theory fits best with her intuitions in the Gettier Cases (step 2). The JTB theory, the causal theory, the defeasibility account, and the reliability account are alternative theories Sophie has to consider (step 3).

Step 4. *The relevant background theories*

⁷See Lehrer & Paxson: A defeasibility condition requires that there is no other true statement, d , such that the conjunction of S 's present evidence for p with d would fail to make S justified in believing p . [72, p. 230].

⁸See, e.g., Goldmann [43].

⁹See, e.g., Goldman [44].

Sophie needs to think about her background theories which will help her to decide between the JTB theory, Clark's account and the alternative theories. Which theory of knowledge Sophie chooses certainly depends on what she thinks justification is. Theories of justification are relevant yet do not directly concern the question of what knowledge is. To keep Sophie's case simple, let us suppose that she has externalist views about justification, which means that she will not consider internalist theories of knowledge. Moreover, metaphysical accounts of causation might influence Sophie's choice. Since metaphysics of causation suggests that abstract and future facts cannot be causes, this counts against the causal theory of knowledge. Other background theories will probably influence Sophie's preferences tacitly.

Step 5. *Restoring coherence*

What Sophie has done so far should help her to remove the inconsistency between the content of her intuition that Smith does not know the relevant proposition and her accepted theory of knowledge, the JTB theory.

Sophie thinks that Clark's account is the one that fits her Gettier intuitions best, but she now wonders whether it covers other cases as well and tries to come up with a counterexample. After a while of thinking, she comes up with a case where the justification does not rest on a false assumption, similar to Keith Lehrer's [71] Nogot case. In Lehrer's case, Nogot in *S*'s office has given *S* evidence that he, Nogot, owns a Ford. *S* directly moves from his evidence to the conclusion that someone in *S*'s office owns a Ford, without arguing via the assumption that Nogot owns a Ford. Nogot does not own a Ford, however, someone else in the office, Havit, owns one, which makes *S*'s belief true. Let us suppose that Sophie's case is very similar to this case in that the subject's reasoning somehow does not rest on a false assumption. Sophie comes to the conclusion that Clark's analysis is not correct, because it cannot account for some cases that are very similar to the original Gettier Cases.

Let us say that the causal theory is not consistent with Sophie's background

theory on causality. The defeasibility analysis she came up with covers more cases than Clark's analysis, but Sophie thinks it is also extremely complicated.

Sophie thinks about the advantages and disadvantages of the JTB theory, the theory she thinks fits her Gettier intuitions best, and the alternative theories. She weighs them against each other and, given her views on justification and causation which she does not want to give up, decides that a reliabilist account of knowledge is the best choice.

While Sophie's case is fictional and could have been told differently, it obviously roughly corresponds to how parts of the debate about the Gettier Cases were conducted in the literature, and some philosophers made a similar decision as Sophie.

6.4 Rational Belief Revision

One main aim of the chapter is to see whether MRE is as trivial as some philosophers have claimed. To specify what exactly they mean, let us look at two obvious ways in which MRE could be interpreted as non-trivial or misguided.

First, some philosophers understand MRE as a theory of justification (e.g., Rawls [98], Daniels [19], [21], [22], Elgin [32], Stich [117], and Goodman [47]). According to MRE as a theory of justification, the beliefs we reach as a result of applying MRE to a certain topic are thereby justified. For instance, if Sophie comes to the conclusion that Smith in the Gettier Case does not know that either Jones owns a Ford or Brown is in Barcelona, then Sophie's belief that Smith does not have knowledge of this proposition is justified if it is in reflective equilibrium, i.e., if it coheres with all her other beliefs about knowledge as a result of having appropriately followed steps 1 to 5. As such, MRE is most naturally understood as a coherentist theory of justification. According to coherentist theories of justification, our beliefs are justified through their

relation to other beliefs.¹⁰ Coherentist theories of justification are controversial, especially for the justification of our beliefs in areas other than moral or political philosophy. Clearly philosophers such as Williamson [140] who claim that reflective equilibrium is trivial do not have reflective equilibrium as an account of justification in mind.

What else could MRE be if not a theory of justification? Besides merely aiming for true beliefs, we aim to build our theories or revise our beliefs in a rational manner. Ideally, we want a methodology in the sense of rules that guide us in the process of building a theory or revising our beliefs. MRE could simply be understood as a method of rational belief revision that does not carry any commitment with regard to what exactly we gain when we apply MRE.¹¹ As a method of rational belief revision, MRE is neutral with respect to epistemic theories such as foundationism or coherentism, and it is neutral with respect to epistemic externalism or internalism as well. Distinguishing a method of rational belief revision from the question whether our beliefs are justified or true accounts for the strong intuition that our opponents sometimes are as rational in holding the beliefs they hold as we ourselves are, even if either our opponents or we ourselves have unjustified and false beliefs.¹² Supposing that we are not in a sceptical situation, MRE as a method of rational belief revision could nevertheless be conducive to the justification of our beliefs. It is *prima facie* plausible that if a subject *S* is not deceived by a Cartesian demon and starts out with beliefs that are rational for her to hold, and if she revises her beliefs in a rational way to reach a theory *T*, then it is more likely that *T*

¹⁰See Daniels [23].

¹¹Philosophers have discussed several ways in which MRE could be a useful method in moral philosophy. Geoffrey Sayre-McCord [103] mentions three approaches that do not involve the justification of beliefs. According to the first, MRE is a *heuristic* method, i.e., MRE is useful to discover the fundamental truths of morality, but it does not justify the beliefs reached through its application. According to a second approach, there is a *moral obligation* to act only upon moral principles that are in reflective equilibrium with all our other beliefs. According to a third and *pragmatic* view, it is in some sense *useful* to act upon a principle which is in reflective equilibrium with all other beliefs we hold. See also Kappel [57], who is pessimistic concerning the role of MRE in epistemic justification, but mentions that a pragmatic or otherwise not truth-related account of MRE might be defensible.

¹²See also Kelly & McGrath [58], who think that MRE ought to lead to beliefs that are rational for us to hold.

is justified than that T is not justified.

In different terms, if S is not victim of a Cartesian demon and starts out with beliefs that are rational for her to hold, and if she revises her beliefs in a rational way to reach a theory T , then this is at least *some* evidence for the truth of T . This does not mean that coherence of T is the only or even the best evidence for the truth of T . However, it could be evidence which is easier to access than the truth of T or some alternative evidence for T , and hence it could be evidence we ought to appeal to in order to establish T (see Chapter 5, where I defend the idea that for methodological reasons, we ought to appeal to evidence that is relatively easy to access). In what follows, I will take MRE to be a method of rational belief revision. I will assume that having applied MRE to a certain subject matter gives us some evidence as to the truth of the resultant theory T on the subject matter.

Here is a second way in which MRE could be interpreted as non-trivial: it could be understood as misguided because it idealizes our practice of revising beliefs. DePaul [25] claims that this is the most serious problem MRE confronts, and that MRE can be understood as idealizing our practice with regard to two aspects. First, the order in which we ought to proceed according to MRE does not correspond with what we in fact do: we hardly ever start out with first determining the relevant intuitions (step 1), figuring out a set of theories that matches (step 2), then taking alternative theories into consideration (step 3), and finally using our background theories (step 4) to build the most comprehensive and coherent theory on the subject matter (step 5). We rather ‘naturally bring all kinds of considerations [...] into play helter skelter as they occur to us’¹³. Second, the quantity of intuitions and theories we take into account is limited. It is simply not possible for us to take *all* relevant alternative theories into account (step 3), and it is not possible to test our intuitions concerning all relevant aspects (step 1), because ‘one would need to reflect upon and form [intuitions] about far too many kinds of hypothetical cases’¹⁴.

¹³DePaul [25, p. xcii].

¹⁴DePaul [25, p. xcii].

Looking at Sophie's case, one might think this is indeed a problem. Surely the way Sophie applies MRE to the Gettier Cases is an idealization. No single philosopher did apply or could have applied MRE in the way Sophie does: Sophie starts out with step 1 and then goes through steps 2 to 5. She also takes far more theories into consideration in steps 2 and 3 before she moves on to the final step 5 than philosophers did in 1963.

There are two ways to reply to the idealization objection. First, we might say that MRE has to be understood as making claims about what we *should ideally do*, not as telling us what we ought to do given our temporal and intellectual constraints. MRE says that we should consider all possible alternatives, but this does not mean that it is not rational for us to stop at a certain point to settle for a coherent theory. Let us look at Sophie again. Even though she considers many alternative theories, one could still criticize her for not doing everything she ought to do in order to appropriately follow the advice MRE gives. Maybe Sophie should not have made a decision as to which theory to endorse, since the debate over the Gettier Cases is ongoing and philosophers are still coming up with theories to cover our intuitions in the Gettier Cases and numerous variations of the Gettier Cases. Ideally, Sophie would even be much smarter. However, just as it seems rational for Sophie to stop considering alternative thought experiments and alternative theories, other philosophers could be rational relative to their limits.¹⁵

It is unproblematic to stop considering alternative theories because MRE can be applied to a minimal set of intuitions and theories. MRE can be re-applied whenever someone comes up with a new theory or a new thought experiment, and the accepted theory can be revised again. This is a reply to both aspects of the idealization objection, the worry that we do not always go through the steps in the right order and the worry that we cannot consider

¹⁵One might want to determine such a point where we could be allowed to stop considering alternative theories. For instance, one could argue that we are entitled to settle for a coherent theory as soon as there is widespread agreement on the subject matter amongst our peers. Or one could argue that we are entitled to stop considering new cases and theories as soon as our arguments are convincing to us. Or one could argue that we are entitled to stop considering new cases and theories as soon as we simply run out of ideas.

all relevant theories and intuitions. To engage in MRE could simply mean to engage in an ongoing series of applications of MRE. As Rawls claims: once a subject has reached a coherent theory,

[...] this equilibrium is not necessarily stable. It is liable to be upset by further examination [...] and by particular cases which may lead us to revise our judgments [...] [99, p. 18]

In fact, as I presented Sophie's case, Sophie comes up with a new thought experiment in step 5, namely with a case that is similar to Lehrer's Nogot case. The reason why I let Sophie come up with this case in step 5 is that she compares different theories in step 5, and hence she thinks about whether Clark's account can accommodate as true all or enough intuitions only in step 5. Strictly speaking, Sophie goes back to step 1. However, it seems that this is not a problem for MRE: the order in which we follow the steps does not seem to be crucial.

A second way to reply to the idealization objection is to say that MRE ought to be applied by a group of researchers instead of a single philosopher. Whereas Rawls, Daniels, and DePaul think that MRE is to be pursued by an individual researcher, Goodman [47] defends such a collective MRE. The view is that as individual philosophers, we work on different ends: we discover inconsistencies, we develop theories that cover our intuitions best, and we develop background theories. We do all this in much detail, which might require a whole career or life-time. As a community of researchers, we might eventually reach the aim of step 5: one single theory wins. Maybe Sophie might better be understood as representing a group of researchers, not an individual philosopher.

DePaul [25] argues that MRE cannot possibly be understood as the description of what we are supposed to do as a group of philosophers:

Because of the way revisions are determined, [reflective equilibrium] must be a first-person inquiry. Propositions do not seem true to a group of people except in the derivative sense that they seem true

to each member of the group. Any disagreement within a group and there will be nothing to determine how conflicts are to be resolved, and hence, nothing to determine the group's state of [reflective equilibrium]. Moral inquiry can be a joint endeavor according to [reflective equilibrium] only insofar as we agree or insofar as one person is willing to assist another in her individual attempt to bring her beliefs into equilibrium by doing such things as pointing out potential conflicts in her beliefs, presenting examples that might elicit interesting intuitions or proposing theories that might account for her [intuitions]. Alternatively, one might approach some other person as a subject, taking that person's beliefs and seemings as data and attempting to work out what that person's state of [reflective equilibrium] would be. [25, p. 1xxx]

DePaul thinks that disagreement in intuitions makes a collective MRE impossible. However, disagreement might simply show that we have not yet reached the final state of reflective equilibrium. There surely is a lot of disagreement in philosophy, and this disagreement concerns our theories as much as our intuitions. We could nevertheless all be concerned with the same project, namely with finding the best theory of a certain subject matter, and it is still possible that in the end one theory will win in the sense that it will be accepted by everyone.

While I do not think that disagreement makes collective MRE impossible, I agree with DePaul that MRE should be understood as a method an individual philosopher ought to pursue. The reason is that if MRE is what a group of philosophers ought to engage in, it is unclear what the advice for the individual philosopher would be and how we would assess whether an individual philosopher is revising her beliefs in a rational way. This, however, is what we are interested in when we talk about a method of rational belief revision. We want such a method to give us advice on how to revise our beliefs as individual philosophers and we want to be able to decide whether an individual

philosopher is revising her beliefs rationally.

Even if other people's intuitions matter, it does not follow that MRE is a collective enterprise. As I argued in Chapter 3, it is plausible to think that we do not only take our own intuitions into consideration but rather rely on other philosophers with respect to intuitions as much as we do with respect to theories. Endorsing MRE as a method for an individual philosopher does not mean that we cannot rely on work other people have done, on their intuitions or on the theories they developed.

If we do not understand MRE as a theory of justification and if we think it does not have to be viewed as an idealization of our practice, the concern is that step 1 to 4 of MRE are trivial and step 5 gives us only very general instructions. According to Rawls, we have to go 'back and forth', sometimes to adjust the principles to our judgments, sometimes to conform the judgments to our principles.¹⁶ Similarly vaguely, Daniels claims that we are 'expected to revise our beliefs at all levels as we work back and forth among them and subject them to various criticisms'.¹⁷ It seems that MRE collapses into a trivial and uncontroversial claim about philosophical methodology, as Timothy Williamson expresses in the following passage about MRE:

The question is not whether philosophers engage in the mutual adjustment of general theory and judgments about specific cases—they manifestly do—but whether such descriptions of it are sufficiently informative for epistemological purposes. [140, p. 244]

Similarly, Foley thinks that MRE is too general to be useful:

The problem with this recommendation is familiar. It is not so much mistaken as unhelpful. At best, it is meta-advice [...] It tells you essentially this: take into account all the data that you think to be relevant and then reflect on the data, solving conflicts in the

¹⁶Rawls [99, p. 20].

¹⁷Daniels [23].

way that you judge best. On the other hand, it does not tell you what kinds of data are relevant, nor does it tell you what is the best way to resolve conflicts among the data. It leaves you to muck about on these questions as best you can. [36, p. 128]

DePaul [24], [25] goes so far as to argue that MRE is the *only rational method* in philosophy. In a nutshell, DePaul's version of MRE says that we should reflect upon the logical and evidential relations between all our relevant beliefs, and that we should resolve conflicts which might emerge during this process in the best possible way we can figure out. DePaul then argues that it is difficult to think of a rational alternative to MRE thus construed, and that an opponent of MRE would have to make one of the following claims: (A) we should abandon reflection altogether; (B) our method should direct the inquirer to reflect, but to do so incompletely, i.e., to leave certain beliefs, principles, or theories out of account; (C) our method should not allow the results of the inquirer's reflections to determine what the inquirer goes on to believe.¹⁸ Unsurprisingly, DePaul concludes from his discussion of (A), (B), and (C) that any method endorsing at least one of these claims would be irrational.

In the recent debate over the methodology of philosophy, different views concerning the role of intuitions in philosophy have been defended. One way of deciding whether MRE is trivial is to see whether it is compatible with these different views. If it is not, then it seems that MRE is not trivial. I argue that MRE is not compatible with the sort of scepticism about the reliability of intuitions advocated by experimental philosophers (section 6.5). I then show that MRE is compatible with different views according to which intuitions (either the psychological states or the contents) play a role in philosophy. However, more specific methodological claims have to be made in order to distinguish the practice I have defended in Chapters 3 to 5 from other views and in order to serve as a methodological guide (section 6.6).

¹⁸DePaul [24, p. 301].

6.5 Reflective Equilibrium and Scepticism

We saw in Chapter 3 that some experimental philosophers draw a radical conclusion from their studies: using intuitions as evidence for or against philosophical theories is an unreliable method which should not be pursued.¹⁹ Some philosophers have claimed that MRE can accommodate scepticism about the reliability of intuitions as advocated by experimental philosophers. I rejected such scepticism, but in order to see whether MRE is trivial, let us look at how it could be compatible with MRE.

I will first look at a view according to which sceptical considerations are not supposed to be part of MRE because meta-theories in general are not supposed to be part of MRE. I will then look at Singer's view which, I will argue, is not compatible with MRE. Finally, I will look at DePaul's view according to which scepticism is compatible with MRE because it could come as a result of applying MRE that we disregard every single intuition. I will argue that if this was the case, MRE would be self-defeating.

MRE and meta-theories

One way to argue that scepticism about the reliability of intuitions is not compatible with MRE is to exclude meta-considerations from the beginning. DePaul mentions that according to Rawls' original account, meta-considerations and arguments from philosophy of language and metaphysics that have been used for or against metaethical theories such as moral relativism, moral realism, or noncognitivism are not supposed to be considered in MRE. The reason is that while MRE is supposed to help us decide between different moral theories, meta-theories do not bear directly on the moral subject matter under consideration.²⁰ The same would apply to contemporary empirical research on the reliability of our intuitions, according to DePaul:

¹⁹See, e.g., Alexander & Weinberg [2], Weinberg [133].

²⁰DePaul [25, p. lxxxix].

Efforts to use results from psychology or neuroscience or evolutionary theory to question the reliability of some or all of our [intuitions] are now extremely prominent [...] As Rawls conceived of [wide reflective equilibrium], these background theories would not be part of the equilibrium. Because they provide premises for a broad skepticism regarding morality, they would not serve as premises of arguments for, or against, any particular [moral theory]. [25, p. lxxxix]

Exactly the same would be true for results from psychology or experimental philosophy in epistemology. Take Sophie and her epistemic intuitions. Sophie is interested to know whether Smith has knowledge in the Gettier Case, and more generally what knowledge is. Questions concerning the relevance of empirical research on intuitions do not bear directly on the question whether Smith has knowledge of the relevant proposition or not, so it seems not to help Sophie to answer her question.

DePaul mentions that Rawls' main reason to exclude meta-theories from MRE is that he is interested in MRE as a method to detect our moral sensibilities rather than to detect the truth about moral issues. In order to determine which moral theory captures our moral sensibilities best, meta-theories are obviously not relevant. To compare it with a simple case: if I am interested to find out about my food preferences and notice that I do not like broccoli, the fact that broccoli is healthy and it would be much better for me to like broccoli is irrelevant to my concern.

However, if we are ultimately interested in the truth of our theories rather than merely in the systematization of our beliefs and intuitions, the question of whether our intuitions are reliable is highly relevant. Hence, if MRE told us to ignore doubts as to their reliability, it could not be a method of rational belief revision. Excluding meta-considerations such as scepticism about the reliability of intuitions from the beginning is therefore not an option for MRE as a method of rational belief revision.

MRE without intuitions

In line with some experimental philosophers, Singer [104] draws sceptical conclusions from recent empirical studies and defends a view according to which we should not assign any plausibility to our moral intuitions. Singer objects to ‘any method of doing ethics that judges a normative theory either entirely, or in part, by the extent to which it matches our moral intuitions’.²¹

As an example of evidence for the insignificance of moral intuitions, Singer discusses Joshua Greene et al.’s [48] studies on intuitions people have when confronted with different versions of the Trolley Case. In one version of the Trolley Case (the one presented in the Introduction to this thesis), we are asked whether a fat man should be pushed down a bridge to stop the trolley, in which case only one person dies and five people who would otherwise be killed survive. Most of us have the intuition that pushing the fat man off the bridge would be wrong. In another version, we are asked if the driver of the trolley should side-track the trolley, which would again kill only one person instead of five people. Most of us have the intuition that the driver should side-track the trolley. Greene et al. [48] conducted brain scans of subjects while they had intuitive reactions to cases very similar to the two versions of the Trolley Case. The results suggest that different intuitive responses have to be explained by differences in the emotional pull of situations which involve causing someone’s death in a close-up, personal way vs. causing someone’s death in a way which is at a distance and less personal.²² Singer thinks that more research is likely to show that Greene ‘has not only explained, but explained away the philosophical puzzle’ of why our intuitions in different versions of the Trolley Case differ. Based on Greene et al.’s results and arguments from evolution²³, Singer draws

²¹Singer [104, p. 346].

²²Greene et. al [48, p. 2106].

²³According to Singer, ‘[...] the salient feature that explains our different intuitive judgments concerning the two cases is that the footbridge case is the kind of situation that was likely to arise during the eons of time over which we were evolving; whereas the standard trolley case describes a way of bringing about someone’s death that has only been possible in the past century or two, a time far too short to have any impact on our inherited patterns of emotional response.’ [104, p. 348].

the following sceptical conclusions for moral intuitions and for MRE:

Thus recent scientific advances in our understanding do have some normative significance, and at different levels. At the particular level of the analysis of moral problems like those posed by trolley cases, a better understanding of the nature of our intuitive responses suggests that there is no point in trying to find moral principles that justify the differing intuitions to which the various cases give rise. Very probably, there is no morally relevant distinction between the cases. At the more general level of method in ethics, this same understanding of how we make moral judgments casts serious doubt on the method of reflective equilibrium. There is little point in constructing a moral theory designed to match considered moral judgments that themselves stem from our evolved responses to the situations in which we and our ancestors lived during the period of our evolution as social mammals, primates, and finally, human beings. [104, p. 348]

We are interested in the second claim on a more general level, according to which empirical studies cast doubt on MRE. Even though Singer does not think it would be a good idea, he mentions that MRE could possibly be interpreted ‘wide enough’ to accommodate a practice where we do not take any of our intuitions into account. In that case, MRE might be compatible with the idea that intuitions should not play any role in philosophy, but it would no longer be a distinctive method for normative ethics.²⁴

Scepticism about the reliability of intuitions entails *prima facie* that step 1 is misguided: we ought not take our intuitions into account. In a concrete case of a thought experiment such as the Gettier Cases, this means that a subject ought not assign any plausibility to the fact that she has an intuition that *P* or to the content of the intuition that *P*:

²⁴Singer [104, p. 347].

No Plausibility. The subject assigns no plausibility to her intuition that P .

A method that rules out intuitions (either the fact that we have an intuition that P or the content of the intuition that P) as evidence from the beginning does not seem to be compatible with MRE, at least not with MRE as presented in section 6.2. According to MRE, we remove an inconsistency between an intuition and our accepted theory either by disregarding the intuition and revising our judgment or by adjusting our accepted theory (or our background theories, or all three). The crucial point is that two options are available with respect to the intuition that P : either to accommodate it as true or to disregard it. In principle, MRE leaves Sophie the choice to either endorse that Smith has no knowledge, or to reject it and either stick with the JTB theory or endorse a different theory. Which option she chooses depends on her background theories and on how good the alternative theories she comes up with are, but neither is ruled out from the beginning.

Proponents of MRE might disagree about how exactly the process of restoring coherence in step 5 ought to take place. We said that background theories are relevant to the areas we are interested in, i.e., the relevant MRE is *wide* MRE. This means that we use arguments from background theories in step 5, and it also means that background theories are amongst the theories that can be adjusted. However, that there are two ways of dealing with the intuition seems to be widely shared. Here is how Ernest Sosa [111] puts it (using ‘principle’ for ‘theory’):

If a conflict pits the intuitive pull of an example against the tug of a familiar principle, we seek to remove or revise one or the other, so as to remove the tension. Sometimes the particular intuition(s) will win, but sometimes the tug of the principle must prevail. [111, p. 262]

Roy Sorensen [110] writes (using ‘theoretical principles’ for ‘theories’ and ‘atheoretical judgment’ for ‘intuition’):

You attain reflective equilibrium when your theoretical principles cohere with your atheoretical judgments. To reach this state, you must remove conflicts between them. Sometimes the conflict is resolved by giving up the principle and sometimes by giving up the intuition. [110, p. 83]

It seems that step 5 of MRE, applied to thought experiments as counterexamples to philosophical theories, entails the following methodological claim with respect to intuitions.

Coherence. In case of a conflict between our intuitions and our accepted theory on a certain subject matter it is sometimes permissible to disregard an intuition that P and sometimes permissible to adjust theory to accommodate as true an intuition that P .

If we endorse No Plausibility, Coherence must be false. Since according to scepticism about the reliability of intuitions, we ought not assign any plausibility to our intuitions, it is not true that we have two options to regain coherence. According to this view, thought experiments simply would not be potential counterexamples and would not have to be taken into consideration at all. Sophie, for instance, would have to disregard her intuition in the Gettier Case and stick with her accepted theory, the JTB theory of knowledge—or use other kinds of evidence to support a different theory of knowledge. According to MRE, however, we have to assign at least *some* initial plausibility to our intuitions. Hence, if scepticism about the reliability of intuitions entails No Plausibility, it is not compatible with MRE.

Scepticism as a consequence of MRE

However, there might be a way to account for scepticism about the reliability of intuitions that does not entail No Plausibility. Let us look at DePaul's view. In contrast to Singer, DePaul thinks that moral intuitions play an important role,

but he agrees that MRE could accommodate scepticism about the reliability of intuitions.²⁵ As a consequence, DePaul's characterization of MRE is indeed so general that it could be true of any kind of truth-directed activity: the essence of MRE is that it 'directs one to leave nothing out of consideration and to believe what seems likely to be true upon due consideration'.²⁶ It seems that MRE understood as broadly as this can no longer count as a method of rational belief revision since it does not entail any advice as to how to revise our beliefs.

DePaul thinks that scepticism about the reliability of intuitions could come as a result of applying MRE. If we take empirical research on particular intuitions such as psychological and neuroscientific evidence into account, it could turn out that we end up rejecting every single one of our intuitions. This seems to be what DePaul has in mind when he says the following:

Suppose the data provided by the research seems much more likely to be true to S than any of her intuitive moral judgments, and that she follows the arguments from the data and has no doubts about them. So, S excises all normative moral beliefs from her overall system. She ends up accepting no moral theory; she makes no moral judgments. She only has beliefs about morality, e.g., that all the moral judgments she previously made were mistaken and that all the moral judgments other people make are mistaken. Does it follow that S would have abandoned the method of [reflective equilibrium]? Not at all—she would have done exactly what that method dictates. [25, p. c]

Since he thinks that all moral theory in the end amounts to intuitions, DePaul concludes that if the subject rejected all her moral intuitions, she would have no moral views at all. We would not have to take this extreme stance on the role of intuitions in non-moral philosophy, but we could imagine ending

²⁵DePaul [25, p. c].

²⁶DePaul [25, p. cii].

up disregarding all intuitions from thought experiments, because it turned out that we have reasons to do so for every single intuition.

DePaul's idea is that since it would come as a result of applying MRE, disregarding all intuitions would be compatible with MRE. However, such a result would certainly raise serious doubts about the reliability of intuitions more generally. We would then have to adjust our meta-philosophical view: it would hardly be rational to consider further intuitions, and we would rather have to assign no initial plausibility to intuitions anymore. We would have to endorse No Plausibility, which entails the denial of Coherence (as discussed above). Hence, if MRE leads to the rejection of every single of our intuitions, it is self-defeating.

However, MRE could be compatible with taking empirical evidence about the reliability of particular intuitions into account, as long as this does not lead us to disregard all our intuitions. In the next section, section 6.6, I discuss how exactly taking such empirical evidence into account is compatible with MRE.

6.6 Reflective Equilibrium and the Appeal to Intuitions

In this section, I show that MRE is compatible with the view that we ought to use the fact that we have an intuition as evidence (as defended in Chapters 3 to 5) as well as with the view that we ought to use the content of an intuition as evidence. I will first look at a view according to which MRE ought to be applied to our psychological states only (e.g., to our beliefs and intuitions about knowledge). I will then look at a view which has it that we ought to apply MRE to the contents of our psychological states (e.g., to the contents of our beliefs and intuitions about knowledge). I will in particular look at the role the fact that we have an intuition could play in MRE according to the second view.

MRE applied to psychological states

As mentioned in section 6.5, MRE has been understood as a method that merely systematizes our opinions. However, that MRE systematizes our beliefs and intuitions does not necessarily mean that it *merely* systematizes our beliefs and intuitions. DePaul thinks that even though a systematization is what MRE aims at, a coherent system of beliefs on a moral subject matter might ‘reveal the plain truth about morality’.²⁷

In his discussion of step 5 of MRE, DePaul claims that it is unclear what our criteria are when we try to work out the most coherent theory, and that proponents of MRE have not paid much attention to this question. He then makes several suggestions, of which at least one seems to clearly concern our psychological states only. DePaul suggests that one might appeal to one’s degree of belief, or to the degree of confidence in the belief, or to the belief’s credence.²⁸ To take our case again, what Sophie would take into account in step 1 and 2 are her current intuitions and beliefs, she would consider possible alternative beliefs in step 3, and she would take her background beliefs into account in step 4. She would then choose the beliefs and intuitions with the highest degree of confidence or with the highest credence in step 5. Hence, she would apply MRE to her psychological states rather than to the contents of her states. According to such an account of MRE, there is no problem when it comes to taking evidence about the reliability of our intuitions into account.

MRE applied to contents

As mentioned in Chapters 3 to 5, Williamson [140] criticizes an approach to philosophical methodology according to which we are concerned with psychological states. Williamson argues that it is not clear how we could get from facts about our psychology to facts about the world such as facts about knowl-

²⁷See DePaul [25, p. xciii], see also DePaul [25, pp. xc-xci].

²⁸DePaul [25, p. lxxx]. DePaul endorses a view according to which one ought to revise one’s beliefs depending on which belief seems to be true to one, see DePaul [25, pp. lxxx-lxxxj].

edge, which are the facts we are ultimately interested in. Having to argue from facts about psychological states to facts about the world provokes scepticism and should therefore be avoided.

Williamson gives two main arguments against MRE understood as a method to be applied to psychological states. First, he argues that there is no point in applying MRE to psychological states when we still rely on some first-order claims: when weighing different beliefs against each other on the basis of criteria such as coherence, we already use first order criteria to decide what is coherent and what is not. We do not use our beliefs about coherence, and we might as well simply talk about the contents of our psychological states.²⁹ Williamson then repeats one of his arguments against the psychologization of evidence, the argument from exceptionalism (see Chapter 4), and applies it to MRE:

To say that mathematicians or biochemists or historians strive to bring their opinions into equilibrium would be sadly inadequate as even a summary description of their method of research [...] Is philosophy so different that in its case such a description will suffice? If so, it should give up any claim to be an evidence-based discipline. [140, p. 246]

In this quote, Williamson presupposes that MRE is meant to be applied to psychological states rather than to the contents of these states. However, note that an interpretation of MRE according to which it has to be applied to our psychological states does not follow in any way from MRE as presented in section 6.1. Daniels' account has it that we use our background theories to argue for and against different views, and we decide on the basis of these arguments which theory to endorse. Such an interpretation of MRE does not presuppose that we are concerned with psychological states only.

If we think that in philosophy, we are concerned with facts about the world (e.g., facts about knowledge) and not merely with facts about our psychological

²⁹Williamson [140, p. 246].

states (e.g., beliefs and intuitions about knowledge), we should understand MRE to be applied to the contents of our psychological states rather than to our psychological states only.

I argued in Chapters 3 to 5 that in our current practice, the fact that we have an intuition plays a role in the process of deciding how to revise our beliefs even if we are interested in facts about the world (e.g., in knowledge) rather than in facts about our having of intuitions (e.g., about knowledge). While the fact that we have an intuition can serve as initial evidence against a philosophical theory, we then move on to further investigations as to why we have the intuition and into the truth or falsity of the content of the intuition.

I moreover showed in section 4.5 of Chapter 4 that we can distinguish several ways of explaining away the significance of an intuition to the truth or falsity of its content. Such an explanation could, e.g., rely on some general psychological dispositions of the intuiters (strategy i)), some pragmatic rules of conversation (strategy ii)), some properties of our reaction to the presentation of the thought experiment scenario (strategy iii)), or some properties of the theory's subject matter (strategy iv)). In all these cases, we give an explanation of the fact that we have an intuition by appeal to some facts other than the truth of the content of the intuition.

In what follows, I will discuss how exactly MRE can account for the appeal to the fact that we have an intuition as evidence. I will consider the practice of appealing to the fact that we have an intuition first as part of step 1 and then as part of step 5 of MRE.

Intuitions in step 1 of MRE

One might think that MRE can accommodate our practice of appealing to the fact that we have an intuition if we take it to be part of step 1. We take the fact that we have an intuition as initial evidence, and then we seek and rule out alternative explanations as to why we have the intuition before we eventually move on with steps 2 to 5 of MRE.

As presented above, in step 1, we select intuitions that are likely to be reliable because we had them under ideal conditions, such as having adequate information about the subject matter and being in a state of mind that is not conducive to error. A view I have endorsed is that only expert intuitions are had under these conditions, so we should only take those intuitions into account. I discussed the so-called expertise defence in Chapter 3, and I described the practice of excluding intuitions on the basis of some general psychological dispositions in Chapter 4 (section 4.5, strategy i)). Take Sophie and the Gettier Cases again. As mentioned in previous chapters, Weinberg, Nichols, & Stich [135] have conducted empirical studies on versions of the Gettier Cases. They presented cases to East Asian, Indian sub-continental and European American undergraduate students. They found that the majority of the European American's intuitions correspond with most philosophers' reaction to the Gettier Cases: the person in the case does not have knowledge. The majority of the East Asian and Indian sub-continental students, however, had different intuitions: they thought that the person in the case has knowledge. Unlike the students who took part in the study, Sophie is not a layperson but an expert epistemologist. If it is correct that philosophers are not prone to the same cultural effects, Sophie should only take her and other expert epistemologists' intuitions into account in step 1 and ignore laypeople's intuitions as irrelevant.

This corresponds to a suggestion Stephen Stich makes. He says that one way in which MRE could be seen as compatible with experimental philosophers' work is

[...] to restrict the class of people whose reflective equilibrium is to count in assessing the justification of inferential principles. For example, Nisbett and I proposed that in saying an inferential principle is justified, what we are saying is that it would pass the (narrow) reflective equilibrium test for those people whom we regard as experts in the relevant inferential domain (Stich and Nisbett 1980). [117, p. 101]

However, not all strategies of explaining away the relevance of an intuition to the truth or falsity of its content can be accommodated in step 1. I showed in Chapter 4 that even if we only take expert intuitions into account, there are still several strategies of explaining away the relevance of an intuition to the truth or falsity of its content (section 4.5, strategies ii) to iv)). Step 1 is not able to accommodate these strategies because they do not serve to define conditions under which our intuitions are reliable. They rather help us to decide whether a particular intuition ought to be trusted and its content accommodated as true in our theory. This, however, is exactly what is supposed to be done in step 5 of MRE. Let us therefore see if there is a way to accommodate these strategies in step 5 of MRE.

Intuitions in step 5 of MRE

We are concerned with the question of whether MRE can accommodate the practice of using to the fact that we have an intuition as initial evidence even if we apply MRE to the contents of our psychological states. If we take the fact that some expert philosophers have an intuition as the relevant initial evidence which gets selected in step 1 of MRE, this fact about our psychological state confronts facts about the subject matter we are interested in. However, facts about our psychology and facts about the subject matter we are interested in are not in direct conflict with each other. Hence, there is no direct conflict to be solved.

To apply it to our case: Sophie is supposed to take the fact that she has an intuition that Smith has no knowledge in step 1 and the JTB theory and alternative theories in step 2 and 3 into account. However, facts about Sophie's psychology and facts about knowledge are not in direct conflict with each other. The fact that Sophie has an intuition about Smith having no knowledge is, moreover, not supposed to be part of her theory of knowledge, just as the fact that she has some other beliefs is not supposed to be part of her theory of knowledge. It is the content of her intuition that Sophie is eventually interested

in, not the fact that she has an intuition.

In other words, given that we are interested in facts about, e.g., knowledge, the problem is that we switch topics when we talk about our intuitions. With respect to the Gödel Case, Jonathan Ichikawa [52] expresses this point as follows:

If I'm to theorize about, say, the nature of reference, I should not feel at all guilty if I fail to explain why people like chocolate, or why the Detroit Lions are so bad. Why should I feel differently about the fact that some people think that in Kripke's story, the name 'Gödel' refers to Schmidt? This psychological fact is interesting, and is, it seems to me, well worth explaining. But it is not clear why it should be the reference theorist's job to explain it. His job is to explain reference, not to explain intuitions about reference. [52, p. 109]

However, if our practice is to first seek and then rule out alternative explanations of our intuitions (as I argued in Chapters 3 to 5), we have to switch to a meta-level before we actually come to our original question.

I do not think that this is a problem. Searching for and ruling out alternative explanations of our intuitions is not the only occasion in which we switch to a meta-level in philosophy. Meta-considerations clearly play a role when we are concerned with philosophical questions such as what knowledge is. Klemens Kappel [57] has suggested a view of MRE that takes theoretical desiderata, i.e., features of our theories, into account in step 5. According to Kappel, many moral theorists as well as moral epistemologists agree that the aim of MRE is to achieve a set of moral beliefs featuring the following epistemic desiderata: consistency, systematicity (a theory should involve explanatory relations), generality (a theory should involve general beliefs that cover a larger area rather than a smaller one) and simplicity (general explanatory beliefs should be few and simple rather than many and complex). Kappel claims that in MRE, in-

creases in consistency, systematicity, generality, or simplicity are purchased at the cost of a decrease in intuitive acceptability.³⁰ Intuitive acceptability is the desideratum according to which sets of beliefs ought to fit our intuitions or, in other words, according to which the contents of our intuitions are made true in our theory. Whether it is true that we trade fit with intuitions against other criteria or not, these criteria certainly play a role when we revise our beliefs. Brian Weatherston [131] discusses the following desiderata for any philosophical theory: the theory shouldn't have too many counterexamples, not too many unacceptable theoretical consequences, and it should be significant and simple.³¹ Even in the sciences, theoretical desiderata such as coherence, simplicity, explanatory power and explanatory unification play an important role when it comes to deciding which theory we ought to endorse.³² If it is correct that it is part of the process of finding the correct theory that we sometimes switch to a meta-level to consider the simplicity, coherence, etc. of our theory, we might as well include meta-considerations concerning the reliability of our intuitions.

One might then wonder why we usually do not also try to rule out psychological, pragmatic or other explanations as to why we hold a certain theory. Nothing in principle speaks against it. If we had a plausible psychological or pragmatic explanation as to why we hold a certain theory, say the JTB theory, this explanation could certainly be taken into account. However, I think there are reasons why we do not give psychological or pragmatic explanations as to why we hold a certain theory. First, our accepted theory usually meets theoretical desiderata such as consistency, systematicity, and generality. Hence, we already have some evidence for the truth of that theory. Second, intuitions from thought experiments as counterexamples to philosophical theories have a special status, because their contents create inconsistencies. It is therefore not surprising that we try to rule out alternative explanations of our intuitions

³⁰Kappel [57, p. 132].

³¹Weatherston [131, pp. 8–10].

³²See, e.g., van Fraassen [125], Kitcher [59], Churchland [17].

before we revise our theory.

MRE can accommodate the practice I have defended in Chapters 3 to 5, but in order to get methodological advice, we have to say something more specific about the role of intuitions. We said that MRE entails the following claim:

Coherence. In case of a conflict between our intuitions and our accepted theory on a certain subject matter it is sometimes permissible to disregard an intuition that P and sometimes permissible to adjust theory to accommodate as true an intuition that P .

The following two methodological claims have to be added, given our current practice as described and defended in Chapters 3 to 5. In the case of a conflict between an accepted theory and an intuition, we ought to

1. search for an explanation as to why we have the intuition that P which is not based on the truth of the content of the intuition and
2. search for the best-fitting theory that can accommodate as true the content of the intuition that P .

MRE can accommodate 1. and 2. But without these further claims, it is not specific enough to give us concrete methodological advice.

6.7 Conclusion

The question I aimed to answer in this chapter was whether MRE understood as a method of rational belief revision reduces to a trivial claim about philosophical methodology. I argued that contrary to what some philosophers think, MRE is not compatible with scepticism about the reliability of intuitions. I showed that MRE is compatible with a view according to which we are concerned with our psychological states only as well as with a view according to which we are concerned with the contents of these states. I moreover showed that MRE is compatible with the practice I defended in Chapters 3 to 5, namely

the practice of appealing to the fact that we have an intuition as initial evidence even if we are interested in the contents of our psychological states.

The conclusion to be drawn is that MRE is not as trivial as some philosophers think because it is not compatible with scepticism about the reliability of intuitions. But as long as we think that either facts about our having of intuitions or the contents of our intuitions play some role in the process of revising our beliefs, our view is compatible with MRE.

In order to get useful methodological advice, however, we need to add more specific methodological claims with respect to intuitions. I suggested such claims in section 6.6. I take it that having followed MRE plus the additional advice can give us some evidence as to the truth of the resultant theory which is relatively easy to access and hence, as I argued in Chapter 5, is evidence we ought to appeal to in order to establish a philosophical theory.

Bibliography

- [1] Ahlström, B. H. *Major Depression and Family Life—The family's way of living with a long-term illness*, <http://hdl.handle.net/2077/19047>. 2009.
- [2] Alexander, J., and Weinberg, J. Analytic epistemology and experimental philosophy. *Philosophy Compass* 2, 1 (2007), 56–80.
- [3] Bach, K. Seemingly semantic intuitions. In *Meaning and Truth*, J. Keim Campbell, M. O'Rourke, and D. Shier, Eds. Seven Bridges Press, New York, 2002, pp. 21–33.
- [4] Bach, K. Applying pragmatics to epistemology. *Philosophical Issues* 18, 1 (2008), 68–88.
- [5] Baker, A. Simplicity. <http://plato.stanford.edu/archives/sum2011/entries/simplicity/> (2011).
- [6] Bealer, G. 'A Priori' knowledge and the scope of philosophy. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 81, 2/3 (1996), 121–142.
- [7] Bealer, G. Intuition and the autonomy of philosophy. In *Rethinking Intuition. The Psychology of Intuition and Its Role in Philosophical Inquiry*, M. R. DePaul and W. Ramsey, Eds. Rowman and Littlefield Publishers, Inc., Lanham, 1998, pp. 201–239.
- [8] Beebe, J. Experimental epistemology. In *Companion to Epistemology*, A. Cullison, Ed. Continuum, forthcoming, pp. 1–19.

- [9] Brown, J. Knowledge and practical reason. *Philosophy Compass* (2008), 1135–1152.
- [10] Brown, J. Thought experiments, intuitions and philosophical evidence. *Dialectica special issue, Justification*. 65, 4 (2011), 493–516.
- [11] Buckwalter, W., and Stich, S. Gender and philosophical intuition. In *Experimental Philosophy*, J. Knobe and S. Nichols, Eds., vol. 2. Oxford University Press, forthcoming.
- [12] Cappelen, H. *Philosophy without Intuitions*. Oxford University Press, Oxford, 2012.
- [13] Carroll, N. *The Philosophy of Horror*. Routledge, New York, 1990.
- [14] Charlton, W. Feeling for the fictitious. *British Journal of Aesthetics* 24 (1984).
- [15] Cherniak, C. *Minimal Rationality*. MIT Press, Cambridge, MA, 1986.
- [16] Christensen, D. *Putting Logic in its Place*. Oxford University Press, New York, 2004.
- [17] Churchland, P. M., and Hooker, C. A., Eds. *Images of Science. Essays on Realism and Empiricism, with a Reply from Bas C. van Fraassen*. The University of Chicago Press, 1985.
- [18] Conee, E. Phenomenal knowledge. *Australasian Journal of Philosophy* 72 (1994), 136–150.
- [19] Daniels, N. Wide reflective equilibrium and theory acceptance in ethics. *The Journal of Philosophy* 76, 5 (1979), 256–282.
- [20] Daniels, N. On some methods of ethics and linguistics. *Philosophical Studies* 37, 1 (1980), 21–36.
- [21] Daniels, N. Reflective equilibrium and archimedean points. *Canadian Journal of Philosophy* 10, 1 (1980), 83–103.

- [22] Daniels, N. *Justice and Justification: Reflective Equilibrium in Theory and Practice*. Cambridge University Press, New York, 1996.
- [23] Daniels, N. Reflective equilibrium. <http://plato.stanford.edu/entries/reflective-equilibrium/> (2003; substantive revision 2011).
- [24] DePaul, M. R. Why bother with reflective equilibrium? In *Rethinking Intuition. The Psychology of Intuition and Its Role in Philosophical Inquiry*. Rowman and Littlefield Publishers, Inc., 1998, pp. 293–309.
- [25] DePaul, M. R. Methodological issues: Reflective equilibrium. In *The Continuum Companion to Ethics*, C. Miller, Ed. Continuum International Publishing Group, London, New York, 2011, pp. Ixxv–cv.
- [26] DeRose, K. Contextualism and knowledge attributions. *Philosophy and Phenomenological Research* 52 (1992), 913–929.
- [27] DeRose, K. Contextualism: An explanation and defense. In *The Blackwell Guide to Epistemology*, J. Greco and E. Sosa, Eds. Blackwell Publishers, 1999.
- [28] Deutsch, M. Experimental philosophy and the theory of reference. *Mind and Language* 24, 4 (2009), 445–466.
- [29] Deutsch, M. Intuitions, counter-examples, and experimental philosophy. *Review of Philosophy and Psychology* 1, 3 (2010), 447–460.
- [30] Devitt, M. Intuitions as evidence. *British Journal for the Philosophy of Science*, 57 (2006), 481–513.
- [31] Earlenbaugh, J., and Molyneux, B. Intuitions are inclinations to believe. *Philosophical Studies* (2009).
- [32] Elgin, C. Z. *Considered Judgment*. Princeton University Press, 1996.
- [33] Evnine, S. J. Believing conjunctions. *Synthese* 118 (1999), 201–227.

- [34] Feagin, S. *Reading with Feeling: The Aesthetics of Appreciation*. Cornell University Press, 1996.
- [35] Feltz, A., and Zarpentine, C. Do you know more when it matters less? *Philosophical Psychology* 23 (2010), 683–706.
- [36] Foley, R. *Working Without a Net: A Study of Egocentric Epistemology*. Oxford University Press, New York, 1993.
- [37] Foot, P. *The Problem of Abortion and the Doctrine of the Double Effect in Virtues and Vices*. Basil Blackwell, Oxford, 1978 (originally in the *Oxford Review*, Number 5, 1967).
- [38] Gendler, T. G., and Kovakovich, K. Genuine rational fictional emotions. In *Contemporary Debates in Aesthetics and the Philosophy of Art*, M. Kieran, Ed. Oxford University Press, Oxford, 2006, pp. 241–53.
- [39] Gendler, T. S. Alief and belief. *Journal of Philosophy* 105, 10 (2008), 634–663.
- [40] Gendler, T. S. Alief in action (and reaction). *Mind and Language* 23, 5 (2008), 552–585.
- [41] Gettier, E. L. Is justified true belief knowledge? *Analysis* 23, 121–123 (1963).
- [42] Gigerenzer, G. *Gut Feelings: The Intelligence of the Unconscious*. Viking Adult, 2007.
- [43] Goldman, A. A causal theory of knowledge. *Journal of Philosophy* 64, 12 (1967), 357/372.
- [44] Goldman, A. What is justified belief? In *Justification and Knowledge: New Studies in Epistemology*. Reidel, Boston, Dordrecht, London, 1979, pp. 1–23.

- [45] Goldman, A. Philosophical intuitions: Their target, their source, and their epistemic status. *Grazer Philosophische Studien* 74 (2007), 1–26.
- [46] Goldman, A., and Pust, J. Philosophical theory and intuitional psycho-evidence. In *Rethinking Intuition. The Psychology of Intuition and Its Role in Philosophical Inquiry*, M. R. DePaul and W. Ramsey, Eds. Rowman and Littlefield Publishers, Inc., 1998, pp. 179–200.
- [47] Goodman, N. *Fact, Fiction, and Forecast*. Harvard University Press, Cambridge, MA, 1955.
- [48] Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., and Cohen, J. D. An fMRI investigation of emotional engagement in moral judgment. *Science* 293 (2001), 2105–2108.
- [49] Gregory, D. Conceivability and apparent possibility. In *Modality: Metaphysics, Logic, and Epistemology*, B. Hale and A. Hoffman, Eds. Oxford University Press, Oxford, 2010, pp. 319–341.
- [50] Grice, H. P. *Studies in the Way of Words*. Harvard University Press, Cambridge, MA/London, 1989.
- [51] Hawthorne, J. *Knowledge and Lotteries*. Oxford University Press, New York, 2004.
- [52] Ichikawa, J. Explaining away intuitions. *Studia Philosophica Estonica* 2.2 (2009), 94–116.
- [53] Ichikawa, J., and Jarvis, B. Thought-experiment intuitions and truth in fiction. *Philosophical Studies* 142, 2 (2007), 221–246.
- [54] Jackson, F. Epiphenomenal qualia. *Philosophical Quarterly* 32 (1982), 127–136.
- [55] Jackson, F. *Postscript on ‘What Mary didn’t know’*. Routledge, London, 1995, pp. 184–189.

- [56] Jackson, F. *From Metaphysics to Ethics*. Oxford University Press, 1997.
- [57] Kappel, K. The meta-justification of reflective equilibrium. *Ethical Theory and Practice* 9 (2006), 131–147.
- [58] Kelly, T., and McGrath, S. Is reflective equilibrium enough? *Philosophical Perspectives Volume 24*, 1 (2010), 325–359.
- [59] Kitcher, P. Explanatory unification. *Philosophy of Science* 48 (1981), 507–531.
- [60] Knobe, J. Intentional action and side effects in ordinary language. *Analysis* 63 (2003), 190–193.
- [61] Knobe, J. Experimental philosophy. *Philosophy Compass* 2, 1 (2007), 81–92.
- [62] Knobe, J., and Nichols, S. An experimental philosophy manifesto. In *Experimental Philosophy*, J. Knobe and S. Nichols, Eds. Oxford University Press, 2008.
- [63] Kornblith, H. The unattainability of coherence. In *The Current State of the Coherence Theory*, J. Bender, Ed. Kluwer, Dordrecht, 1989.
- [64] Kornblith, H. Appeals to intuition and the ambitions of epistemology. In *Epistemology Futures*, S. Hetherington, Ed. Clarendon Press, Oxford, 2006, pp. 10–25.
- [65] Kornblith, H. Naturalism and intuitions. *Grazer Philosophische Studien* 74 (2007), 27–49.
- [66] Kripke, S. *Naming and Necessity*. Harvard University Press, Cambridge, MA, 1972/1980.
- [67] Krippendorff, K. *Content Analysis. An Introduction to its Methodology*, 2 ed. Sage Publications Inc., Thousand Oaks, CA, 2004.

- [68] Kyburg, H. Conjunctivitis. In *Induction, Acceptance, and Rational Belief*, M. Swain, Ed. Reidel, Dordrecht, 1970.
- [69] Lackey, J. Acting on knowledge. *Philosophical Perspectives* 24 (2010), 361–382.
- [70] Lackey, J. Assertion and isolated secondhand knowledge. In *Assertion*, J. Brown and H. Cappelen, Eds. Oxford University Press, 2011, pp. 251–275.
- [71] Lehrer, K. Knowledge, truth, and evidence. *Analysis* 25 (1965), 168–175.
- [72] Lehrer, K., and Paxson, T. Knowledge: Undefeated justified true belief. *Journal of Philosophy* 66 (1969), 225–237.
- [73] Lewis, D. *Counterfactuals*. Basil Blackwell, Oxford, 1973.
- [74] Lewis, D. *Philosophical Papers*, vol. 1. Oxford University Press, Oxford, 1983, ch. Postscript to ‘Mad Pain and Martian Pain’.
- [75] Lewis, D. What experience teaches. *Proceedings of the Russellian Society* 13 (1988), 29–57.
- [76] Lindseth, A., and Norberg, A. A phenomenological hermeneutical method for researching lived experience. *Scandinavian Journal of Caring Sciences* 18, 2 (2004), 145–153.
- [77] Lowe, E. J. What is the ‘problem of induction’? *Philosophy* 62, 241 (1987), 325–340.
- [78] Ludwig, K. The epistemology of thought experiments: First person versus third person approaches. *Midwest Studies in Philosophy* 31, 1 (2007), 128–159.
- [79] Lycan, W. Moral facts and moral knowledge. In *Judgment and Justification*, W. Lycan, Ed. Cambridge University Press, Cambridge, 1988.

- [80] Lycan, W. G. *Real Conditionals*. Oxford University Press, Oxford, 2001.
- [81] Machery, E., Mallon, R., Nichols, S., and Stich, S. Semantics, cross-cultural style. *Cognition* 92 (2004), B1–B12.
- [82] Machery, E., Olivola, C., and de Blanc, M. Linguistic and metalinguistic intuitions in the philosophy of language. *Analysis* 69 (2009), 689–694.
- [83] Matthews, R. J. Measurement-theoretic accounts of propositional attitude. *Philosophical Compass* (2011), 1–14.
- [84] McGee, V. A counterexample to modus ponens. *Journal of Philosophy* 82 (1985), 462–471.
- [85] Nadelhoffer, T., and Nahimas, E. The past and future of experimental philosophy. *Philosophical Explorations* 10, 2 (2007), 123–149.
- [86] Nagel, J. Intuitions and experiments: A defense of the case method in epistemology. *Philosophy and Phenomenological Research* (forthcoming).
- [87] Nagel, J. Mindreading in gettier cases and skeptical pressure cases. In *Knowledge Ascriptions*, J. Brown and M. Gerken, Eds. Oxford University Press, forthcoming.
- [88] Neta, R., and Phelan, M. Evidence that stakes don't matter for evidence. Available from http://experimentalphilosophy.typepad.com/experimental_philosophy/2008/07/evidence-that-s.html.
- [89] Nichols, S., and Knobe, J. Moral responsibility and determinism: The cognitive science of folk intuitions. *Nous* 41 (2007), 663–685.
- [90] Nolan, D. Defending a possible-worlds account of indicative conditionals. *Philosophical Studies* 116, 3 (2003), 215–269.
- [91] Olweus, D. *The revised Olweus bully victim questionnaire*. University of Bergen, Bergen, Norway, 1996.

- [92] Plantinga, A. *Warrant and Proper Function*. Oxford University Press, 1993.
- [93] Priest, G. Contradiction, belief and rationality. In *Proceedings of the Aristotelian Society* (1985/1986), no. 86, pp. 99–116.
- [94] Pritchard, D. *Epistemic Luck*. Oxford University Press, 2005.
- [95] Pryor, J. The sceptic and the dogmatist. *Noûs* 34, 4 (2000), 517–549.
- [96] Pust, J. *Intuitions as Evidence*. Garland Publishing, New York/London, 2000.
- [97] Radford, C. How can we be moved by the fate of Anna Karenina? In *Proceedings of the Aristotelian Society, Supplemental Vol.* (1975), no. 49, pp. 67–80.
- [98] Rawls, J. Outline of a decision procedure for ethics. *The Philosophical Review* 60, 2 (1951), 177–197.
- [99] Rawls, J. *A Theory of Justice*. Cambridge University Press, Harvard, 1971.
- [100] Rawls, J. The independence of moral theory. *Proceedings and Addresses of the American Philosophical Association XLVII* (1974/1975), 5–22.
- [101] Resnik, M. D. Normative or descriptive? The ethics of belief or a branch of psychology? *Philosophy of Science* 52, 2 (1985), 221–238.
- [102] Roush, S. *Tracking Truth. Knowledge, Evidence, and Science*. Clarendon Press, Oxford, 2005.
- [103] Sayre-McCord, G. Coherentist epistemology and moral theory. In *Moral Knowledge?*, W. S.-A. . M. Timmons, Ed. Oxford University Press, 1996, pp. 137–189.
- [104] Singer, P. Ethics and intuitions. *The Journal of Ethics* (2005), 331–352.

- [105] Sinnott-Armstrong, W., Moor, J., and Fogelin, R. A defense of modus ponens. *Journal of Philosophy* 83, 5 (1986), 296–300.
- [106] Skowronski, D., Serres, G. D., Crowcroft, N., Janjua, N., and Boulianne, N. Association between the 2008-09 seasonal influenza vaccine and pandemic H1N1 illness during spring-summer 2009: Four observational studies from Canada. *PLoS Med* 7, 4 (2010).
- [107] Smart, J. Ockham's razor. In *Principles of Philosophical Reasoning*, J. Fetzer, Ed. Rowman & Allanheld, Totowa, NJ, 1984, pp. 118–128.
- [108] Smuts, A. The paradox of painful art. *Journal of Aesthetic Education* 41, 3 (2007), 59–77.
- [109] Sober, E. What is the problem of simplicity? In *Simplicity, Inference and Modelling: Keeping It Sophisticatedly Simple*, A. Zellner, H. Keuzenkamp, and M. McAleer, Eds. Cambridge University Press, Cambridge, 2001, pp. 13–31.
- [110] Sorensen, R. *Thought Experiments*. Oxford University Press, New York, 1992.
- [111] Sosa, E. Equilibrium in coherence? In *The Current State of the Coherence Theory*, J. Bender, Ed. Kluwer: Academic Press, 1989, pp. 242–250.
- [112] Sosa, E. Minimal intuition. In *Rethinking Intuition. The Psychology of Intuition and Its Role in Philosophical Inquiry*, M. R. DePaul and W. Ramsey, Eds. Rowman and Littlefield Publishers, Inc., Lanham, 1998, pp. 257–269.
- [113] Sosa, E. Skepticism and contextualism. *Philosophical Issues* 10 (2000).
- [114] Sosa, E. Intuitions and truth. In *Truth and Realism*, P. Greenough and M. Lynch, Eds. Clarendon Press, Oxford, 2006.
- [115] Sosa, E. *A Virtue Epistemology*. Oxford University Press, Oxford, 2007.

- [116] Stanley, J. *Knowledge and Practical Interests*. Oxford University Press, 2005.
- [117] Stich, S. Reflective equilibrium, analytic epistemology and the problem of cognitive diversity. *Synthese* 74, 3 (1988).
- [118] Sturgeon, S. Reason and the grain of belief. *Noûs* 42, 1 (2008), 139–165.
- [119] Swain, S., Joshua, A., and Weinberg, J. The instability of philosophical intuitions: Running hot and cold on truetemp. *Philosophy and Phenomenological Research* LXXVI, 1 (2008).
- [120] Thomson, J. J. Killing, letting die, and the trolley problem. *The Monist* (1976), 204–217.
- [121] Thomson, J. J. The trolley problem. *Yale Law Journal* 94 (1985), 1395–1415.
- [122] Tolhurst, W. Seemings. *American Philosophical Quarterly* 35 (1998), 293–302.
- [123] Tye, M. *Consciousness Revisited*. MIT Press, Cambridge, MA, 2009.
- [124] Unger, P. An analysis of factual knowledge. *Journal of Philosophy* 65, 6 (1968), 157–170.
- [125] van Fraassen, B. *The Scientific Image*. Clarendon Press, Oxford, 1980.
- [126] van Inwagen, P. Materialism and the psychological-continuity account of personal identity. In *Philosophical Perspectives*, J. Tomberlin, Ed., vol. 11. Blackwell Publishers, 1997.
- [127] Veenstra, R., Lindenberg, S., Munniksmma, A., and Dijkstra, J. K. The complex relation between bullying, victimization, acceptance, and rejection: Giving special attention to status, affection, and sex differences. *Child Development* 81, 2 (March/April 2011), 480–486.

- [128] Vogel, J. Internalist responses to scepticism. In *The Oxford Handbook of Scepticism*, J. Greco, Ed. Oxford University Press, Oxford, 2008.
- [129] Walton, K. Fearing fictions. *Journal of Philosophy* 75.1 (1978), 5–27.
- [130] Walton, K. *Mimesis as Make-Believe*. Harvard University Press, Cambridge, MA, 1990.
- [131] Weatherson, B. What good are counterexamples. *Philosophical Studies* 115 (2003), 1–31.
- [132] Weatherson, B. Evidence neutrality. Available from <http://brian.weatherson.org>.
- [133] Weinberg, J. How to challenge intuitions empirically without risking skepticism. *Midwest Studies in Philosophy* 31 (2007), 318–343.
- [134] Weinberg, J., Gonnerman, C., Bruckner, C., and Alexander, J. Are philosophers expert intuiters? *Philosophical Psychology* 23, 3 (2010), 331–355.
- [135] Weinberg, J., Nichols, S., and Stich, S. Normativity and epistemic intuitions. *Philosophical Topics* 29 (2001), 429–460.
- [136] Williamson, T. *Vagueness*. Routledge, London, 1994.
- [137] Williamson, T. *Knowledge and its Limits*. Oxford University Press, Oxford, 2000.
- [138] Williamson, T. Philosophical ‘intuitions’ and scepticism about judgement. *Dialectica* 58, 1 (2004), 109–153.
- [139] Williamson, T. Armchair philosophy, metaphysical modality and counterfactual thinking. *Proceedings of the Aristotelian Society* 105 (2005), 1–23.
- [140] Williamson, T. *The Philosophy of Philosophy*. Blackwell Publishing Ltd, 2007.

- [141] Williamson, T. Philosophical expertise and the burden of proof. *Metaphilosophy* 42, 3 (2011), 215–229.
- [142] Wright, C. On epistemic entitlement: Warrant for nothing (and foundations for free?). *Aristotelian Society Supplementary Volume*, LXXVIII (2004), 167–212.
- [143] Zellner, A., Keuzenkamp, H. A., and McAleer, M. *Simplicity, Inference, and Modelling*. Cambridge University Press, 2001.