The Method in Political Philosophy and the Challenge from Experimental Philosophy

Sebastian Johansen Conte
# Table of Contents

Acknowledgments .................................................................................................................... v
Summary .................................................................................................................................... ix
List of Abbreviations ................................................................................................................ xi

Introduction ...................................................................................................................................... 1

PART 1: BACKGROUND AND GROUNDWORK FOR THE THESIS ........................................... 11

1 Overview of Part 1 .................................................................................................................. 13

2 Moral Intuitions and the Method in Moral Inquiry ............................................................... 16
    2.1 Moral Intuitions ............................................................................................................ 16
    2.2 The Role of Intuition in Moral Inquiry ......................................................................... 20
    2.3 The Method of Reflective Equilibrium ......................................................................... 22
    2.4 Taking Stock ................................................................................................................ 27

3 The Challenge from Experimental Philosophy ................................................................. 29
    3.1 Experimental Philosophy ........................................................................................... 29
    3.2 Arguments from Experimental Philosophy: The Argument Scheme .......................... 29
    3.3 Arguments from Experimental Philosophy: A Note on the First Premise .................. 30
    3.4 Arguments from Experimental Philosophy: A Note on the Second Premise .............. 32
    3.5 Is Experimental Philosophy a Game-Changer? ......................................................... 33
    3.6 Challenging the Method in Moral Philosophy ........................................................... 35
    3.7 A Note on the Controversy of Experimental Arguments ........................................... 36

4 Questioning the Role of Intuitions in Philosophical Inquiry ............................................. 39
    4.1 Intuition-Terminology ............................................................................................... 40
    4.2 Explaining Salient Facts about the Philosophical Practice ........................................ 43
    4.3 Case Studies ................................................................................................................. 44
    4.4 Taking Stock ................................................................................................................ 48

5 Picking Up the Threads and Outlining the Contribution of the Thesis .............................. 49

6 Groundwork for the Thesis ................................................................................................. 52
    6.1 Ethical Considerations ............................................................................................... 52
    6.2 Relationship to the Philosophy of Science ............................................................... 53

7 Presentation of the Articles .................................................................................................. 69
TABLE OF CONTENTS

7.1 How the Articles Relate to Each Other ............................................................... 69
7.2 First Article ....................................................................................................... 70
7.3 Second Article .................................................................................................. 71
7.4 Third Article ..................................................................................................... 72
7.5 Suggestions for Further Research .................................................................. 74

PART 2: ARTICLES ..................................................................................................... 77

8 Are Intuitions Treated as Evidence? Cases from Political Philosophy .......... 79
8.1 Introduction ...................................................................................................... 79
8.2 Centrality, Intuitions, and Cases ..................................................................... 81
8.3 The Intuitions Deniers and the Positive-Evidence Challenge ....................... 82
8.4 Analytical Framework ..................................................................................... 83
8.5 The Method of Cases and Expensive Tastes .................................................... 88
8.6 Objections and Replies .................................................................................... 98
8.7 Conclusion ........................................................................................................ 101

9 The Trolley Problem and Intuitional Evidence ............................................... 103
9.1 Introduction ..................................................................................................... 103
9.2 Backdrop ........................................................................................................ 104
9.3 Cappelen and the Trolley Problem ................................................................. 105
9.4 Thomson on Her Method ............................................................................... 108
9.5 A Reanalysis of Thomson’s Article ................................................................. 109
9.6 Explaining Why Cappelen is Wrong ............................................................... 117
9.7 Objections and Replies .................................................................................... 118
9.8 Conclusion ........................................................................................................ 122

10 The Reflective Equilibrium and the Challenge from Experimental Philosophy .... 123
10.1 Introduction ................................................................................................... 123
10.2 The Method of Reflective Equilibrium and Intuitions as Data ................. 125
10.3 Experimental Philosophy and Moral Intuitions ........................................ 130
10.4 Narrow MRE and the Challenge from Experimental Philosophy ............ 132
10.5 Wide MRE and the Challenge from Experimental Philosophy ............... 133
10.6 Is MRE Vacuous? ........................................................................................ 138
10.7 Practical Implementations of MRE and Experimental Philosophy ........... 140
10.8 Conclusion ..................................................................................................... 141

References ............................................................................................................. 143
Acknowledgments

I want to thank my supervisors Robert Huseby and Jakob Elster, who have provided me with excellent comments on various drafts and engaged in lengthy discussions concerning my project. The feedback and support I have received from Robert and Jakob have especially helped me develop my ideas and arguments, and more generally, have been crucial for my development as a scholar. I have learned a great deal from Robert and Jakob, and I am extremely grateful for their support, both on an academic and personal level, throughout these past years.

I also want to thank Herman Cappelen, who supervised me during the first year of the project. My discussions with Herman have been invaluable for gaining a better understanding of the literature on the role of intuitions in philosophy and in developing my arguments.

I am also grateful to my employer, the Department of Political Science (ISV), University of Oslo, which has provided me with generous funding to undertake my project. This made possible for me to go to several workshops and conferences to present my work. Moreover, I am grateful to the department for taking several measures to alleviate any negative impacts of the COVID-19 pandemic on my project. More generally, the department has provided us with an excellent working environment and outstanding colleagues.

The department administration has assisted me with various tasks, such as organizing my teaching and the educational component of the PhD programme, for which I am very grateful. A special thanks to Ilze Gehe, from whom I have learned a lot about research administration and for always helping me with whatever confused me at any particular point in time.

The group of PhD candidates, post-doctoral research fellows, and researchers I have worked with have been an amazing group. A special thanks to Erlend Langørgen, Sigurd Lindstad, Leif Erlend Kårtvedt, Nina Maureen Cadorin, Ingvild Reymert, and Karin Fossheim, fellow PhD candidates and colleagues for most of my time as a PhD candidate at ISV. Having colleagues with whom to share the ups and downs of my PhD project helped tremendously when push came to shove and generally made ISV a very enjoyable environment to work within.

During my project, I have had the privilege of being part of the political theory group at ISV. It has been stimulating to discuss topics in political theory with this group of people, and the reading group organized by Sigurd has been very enjoyable. Engaging with this group has been instrumental in my development as a scholar, and I am grateful to all those individuals who have been part of this group during my time at ISV: Robert, Sigurd, Kim Angell, Raino Malnes, Jonathan Kuyper and Fredrik Dybfest Hjorthen.
ACKNOWLEDGMENTS

In early 2020, I was fortunate to visit Yale University. I would like to thank Joshua Knobe for being such a gracious and supportive host. I would also like to thank the participants of the lab meetings organized by Josh for the interesting presentations and discussions, as well as for the very helpful comments on a presentation of mine. I would also like to thank my companions in New Haven for making my time there so enjoyable.

My stay at Yale University would not have been possible without funding from Fulbright and Wedel Jarlsberg's Fond, for which I am most grateful. Moreover, I extend my thanks to the Norwegian Fulbright Office for helping with the administrative side of my visit to the United States, and, in particular, the assistance I received during the pandemic. A special thanks to Cathrine Schrumpf Nordahl and Petter Næss.

As a part of the Fulbright program, I was in touch with Andreas Føllesdal. I am grateful to Andreas for showing interest in my project and providing me with very useful input.

In 2020/2021, I was a participant in the project entitled What is a Good Policy? Political Morality, Feasibility, and Democracy (GOODPOL). The project was hosted by the Center for Advanced Study (CAS). I am grateful to the principal investigators of the project, Cathrine Holst and Jakob, for inviting me to be a participant in the project. I learned a great deal during my time at CAS and from the workshops and seminars organized as a part of the project. I want to thank all the members of GOODPOL and the people at CAS.

The different parts of this thesis have benefitted greatly from comments from and discussion with a great many people. Thanks to all who have contributed! With regard to Part 1 of the thesis, a previous version of the thesis was presented and defended at a trial public defense in which Klemens Kappel and Jon Hovi served as my opponents. I am grateful to Klemens and Jon for their comments on and discussions of thesis. Part 1 was also presented at a PhD seminar at ISV, and my thanks to Erlend, Tor Gaute Syrstad, Tamta Gelashvili, Arno van Hootegem, and Anna Nordnes Helgøy for their helpful comments. Part 1, and more specifically, the section on the philosophy of science, benefitted from a discussion with Raino, whom I also thank. Furthermore, as always, Robert and Jakob have read and commented on several versions of Part 1, which I greatly appreciated. (With regard to the articles, please see the acknowledgements in Part 2.)

I am truly grateful to my friends and family for their steadfast support. It would have been much tougher to complete this project without them by my side. Having all of you to lean on during the process has been invaluable. A special thanks to Sivert Holth Hartting, Sabina Tica, and Tor for their support and for helpful discussions of my project. The last few months of the
ACKNOWLEDGMENTS

project were particularly stressful, and I especially want to thank Sara Bjønness Hagfors for regularly checking in and lifting my spirits during this final push.

Finally, I would like to thank my partner, Inés Moldskred-Belli. My endless gratitude to Inés for sticking with me through thick and thin during this entire process.

Sebastian J. Conte
Oslo, May 2022
Summary

An important methodological issue in political philosophy (and moral philosophy, more generally) is how to justify moral and political principles and theories. For example, how should we go about justifying a particular articulation of the political ideal of equality? It is a common view that intuitions about cases play an important role in justifying such theories and principles. Although quite commonplace, the view that such intuitions play an important epistemic role in moral inquiries has nevertheless had its fair share of critics for some time. Recently, with the rise of experimental philosophy, new efforts have been made to challenge this view of the role of intuitions. Experimental philosophers employ findings and methods mainly from psychology, to study intuitions about cases. Experimental philosophy studies, it is argued, have revealed that intuitions have properties that suggest that they should not play an important epistemic role. Moreover, some have suggested that this, in turn, suggests that the current mainstream method in political philosophy, the method of reflective equilibrium (MRE), should be rejected. I refer to this argument against MRE as “the challenge from experimental philosophy.” The idea underlying the challenge seems to be that intuitions about cases play an essential role in MRE. The aim of this thesis is to investigate whether the challenge from experimental philosophy succeeds in undermining MRE.

In my investigation of the challenge from experimental philosophy, I grant, for the sake of argument, that when we consider the empirical evidence, the empirical premise of the argument from experimental philosophy is plausible. That is, I grant that the currently available empirical evidence suggests that intuitions have the properties that experimental philosophers claim they have. I assume, then, that when considered in the light of the currently available empirical evidence, the empirical assertions about intuitions are fine. Moreover, in discussing the challenge, I am primarily interested in investigating the assumption that the criterion of intuitive fit is an essential feature of MRE. Therefore, I will for part of the discussion grant the argument from experimental philosophy (and not only that the empirical premise is supported by the currently available evidence).

Furthermore, I assume that most philosophers apply MRE. Finally, I distinguish between two interpretations of the challenge. According to the first interpretation, the target of the challenge is MRE in its ideal form. According to the second interpretation, the target of the challenge is the way MRE is currently applied.

In investigating the challenge from experimental philosophy, I first consider an intriguing response to the challenge from experimental philosophy. According to this response, contrary
to what is commonly believed, intuitions do not play a significant epistemic role in philosophical inquiries. If most philosophers apply MRE, but do not rely on intuitions, then the challenge from experimental philosophy does not undermine MRE. I argue, however, that intuitions play a significant epistemic role in moral inquiries (Articles 1 and 2).

Second, I consider whether intuitions play an essential epistemic role in MRE (Article 3). It could, for example, be the case that granting intuitions an important epistemic role is merely optional. I argue that intuitions do not play an essential role in (some) wide interpretations of the ideal of MRE. Therefore, the challenge from experimental philosophy does not undermine those versions of MRE. However, I also argue that the challenge does undermine practical implementations of MRE according to which the assumption that intuitions play a central epistemic role is treated as fixed. The arguments in Articles 1 and 2 suggest that such implementations are quite common. Thus, since most philosophers apply MRE and treat intuitions as playing an important epistemic role, the challenge from experimental philosophy puts pressure on the current implementations of MRE.
List of Abbreviations

CMJ considered moral judgment
EAA equal access to advantage
EOW equality of opportunity for welfare
ER equality of resources
EW equality of welfare
JTB justified true belief
MRE the method of reflective equilibrium
NESH The National Committee for Research Ethics in the Social Sciences and the Humanities
RE reflective equilibrium
Introduction

Roughly, the study of politics can be separated into political science and political philosophy (sometimes referred to as normative political theory) (see List and Valentini 2016, 526). The aim of political science is mainly descriptive, that is, to describe, explain, and predict political phenomena. For example, political science is concerned with describing and comparing political systems in different countries and investigating what causes wars. Political philosophy is concerned with evaluative and moral questions. Such questions encompass questions such as “What makes a society just?”

There are methodological questions concerning both descriptive and normative inquiries into politics. In general, methodological issues are issues concerning how we should go about investigating our research questions (Box-Steffensmeier, Brady, and Collier 2008). For example, researchers do not have direct access to the content of people’s mental states. Thus, one important methodological issue in political science is how to study people’s mental states. This issue is particularly important for political scientists doing research on people’s attitudes. How to justify political principles is an important methodological issue in political philosophy. This thesis addresses methodological issues in contemporary analytical political philosophy. More specifically, it is concerned with the role intuitions play in justifying theories. Before outlining the specific research questions of the thesis, I first provide some more background information.

---

1 That there is such a distinction between normative and descriptive inquiries of politics is not to say, for example, that scientists engaged in descriptive political science never reveal their attitude towards the subject. Note, however, that the scientist engaged in descriptive political science’s main concern, contrary to a theorist engaged in political philosophy, is not to justify her judgment concerning the value of something. See also Gerring and Joshua Yesnowitz (2006) and List and Valentini (2016, 526-527) for discussion of the relation between political science and political philosophy.
It is a common view among philosophers that intuitions play an important role in inquiries in political philosophy (Floyd 2017a) and moral philosophy (Bedke 2010; McMahan 2013, 106), and philosophy more generally (Bealer 1996; Cath 2012; Malmgren 2011; Pust 2016). More specifically, the idea is that intuitions confer some positive epistemic status and therefore that one criterion for the adequacy of a moral or philosophical theory is that it achieves an overall fit with our intuitions. Following Nikil Murkerji (2014, 229), we may refer to this theoretical desideratum as “the criterion of intuitive fit.” A theory in tune with more of our intuitions is more plausible than a theory in tune with fewer of our intuitions, other things being equal. To illustrate, intuitively, one should help a small child drowning in a pond even if this would ruin one’s new shoes. Thus, a theory consistent with this intuition is a more plausible theory than a theory inconsistent with the intuition, other things being equal.

Although quite commonplace, this criterion has for some time had its fair share of critics (see e.g. Singer 1974; Unger 1996). For example, one worry has been that our intuitions are the product of prejudice. However, such worries were not backed up empirical testing of those specific claims. Lately, however, with the rise of experimental philosophy, new efforts have been made to challenge the criterion of intuitive fit. Experimental philosophers employ findings and methods mainly from psychology, including cognitive neuroscience, to study intuitions (Alfano, Loeb, and Plakias 2018; Fischer and Collins 2015; Mallon 2016). The findings of experimental philosophy studies have been used to challenge the criterion of intuitive fit. To be more specific, the findings of experimental studies are assumed to provide us with descriptive information about intuitions, that is, that they have a certain property. This information is then combined with an epistemic assumption stating that the discovered property suggests that intuitions should not be treated as conferring any positive epistemic status (Mallon 2016). From these premises it is inferred that the criterion of intuitive fit is flawed. Thus, it is

---

2 It is worth commenting on the relationship between moral and political philosophy. In this thesis, I view political philosophy as a sub-discipline of moral philosophy. With regard to the political moralism versus realism debate, then, this view of political philosophy amounts to accepting moralism. Roughly, realism is the view that there is a distinctive political normativity (Rossi and Sleat 2014). In accordance with moralism, I am not assuming that there is a distinct political normativity, but that the issues that political philosophy is concerned with is at the bottom a sub-class of moral issues. Thus, I will often speak of inquiries conducted by political philosophers simply as moral inquiries.

3 This case is from Singer (1972).

4 See also Huemer (2005) for an overview of both older and more recent objections to the criterion of intuitive fit. Note, however, that Huemer discussed the criterion of intuitive fit in relation to a particular metaethical position—i.e. moral realism.

5 Although experimental philosophy is coined “experimental,” not all experimental philosophy studies are experiments. Experimental philosophers apply several different approaches, such as survey studies and corpora studies.
only when findings from experimental philosophy studies are combined with an epistemic assumption that the findings shed light on the epistemic credentials of intuitions (see e.g. Sauer 2018, 218). I will refer to such arguments as “arguments from experimental philosophy.”

A caveat is in order. So far, experimental philosophy studies have examined intuitions about cases. Thus, arguments from experimental philosophy target a restricted version of the criterion of intuitive fit according to which the intuitions theories should conform to are intuitions about cases. In the remainder of this thesis, I will have this restricted version of the criterion in mind when referring to “the criterion of intuitive fit.”

Some, however, not only claim that the criterion of intuitive fit is flawed, but also claim that it follows from arguments from experimental philosophy that philosophers have to reject the method they currently apply (see e.g. Paulo 2020; Singer 2005). I will refer to this inference as “the challenge from experimental philosophy.” The underlying idea of this challenge seems to be that the criterion of intuitive fit is a central part of mainstream philosophical method, and hence that if it is rejected as a desideratum, this will deprive proponents of the method of the necessary resources for philosophizing. So although current practice may involve other desiderata in addition to intuitive fit, the rejection of the criterion of intuitive fit, due to its centrality, will be enough to leave the mainstream method inutile and flawed. This, then, distinguishes the challenge from experimental philosophy from arguments from experimental philosophy. The arguments from experimental philosophy (attempt to) undermine the criterion of intuitive fit. However, whether undermining the criterion of intuitive fit poses a challenge to a particular method depends on whether the criterion is supposed to play an essential role in the method. The challenge from experimental philosophy, then, only gets going by asserting that the criterion of intuitive fit is essential to the method it targets.

The aim of this thesis is to investigate the challenge from experimental philosophy. In doing so, I grant, for the sake of argument, that when we consider the currently available empirical evidence, the empirical premise of the argument from experimental philosophy is plausible. That is, I grant that the empirical evidence suggests that intuitions have the properties that experimental philosophers claim they have. For example, I will grant that when we consider the available evidence from fMRI investigations and other relevant investigations (see e.g.

---

6 Selim Berker (2009) argues that in targeting intuitions about cases, arguments from experimental philosophy rely on more general intuitions. The assertion that the discovered property suggests that intuitions should not be treated as conferring any positive epistemic status is itself backed up by general intuitions about the normative relevance of the property (see also Tersman 2012).

7 See also Singer (2005, 348), for a distinction similar to that between arguments from experimental philosophy and the challenge from experimental philosophy.
Greene 2014), the findings suggest that moral intuitions are the output of affective processes shaped by our evolutionary history.\textsuperscript{8} I assume, then, that when considered according to these terms, the empirical assertions about intuitions are fine.\textsuperscript{9}

Moreover, in discussing the challenge, I am primarily interested in investigating the move from the arguments from experimental philosophy to the challenge from experimental philosophy. As far as possibly, I will grant the argument from experimental philosophy (and not only that the empirical premise of such arguments is supported by the currently available evidence). However, looking ahead, some versions of MRE assert that questions such as what the proper data for moral theorizing is and what properties make something suited for serving as data for moral theorizing cannot be decided independently from our inquiries into substantive moral issues. When discussing such versions of MRE it would not make sense to grant the arguments from experimental philosophy since according to such versions of the method, it is by applying MRE such arguments are justified. That said, I discuss whether it would be a problem for such interpretations of the method if the argument from experimental philosophy should prevail in one’s inquiry.

What I want to investigate, then, is whether, when granting these assumptions, the challenge from experimental philosophy succeeds in demonstrating that the current mainstream method in political philosophy must be rejected. To pursue this aim, this thesis will first investigate an intriguing response to the challenge from experimental philosophy, according to which it is simply wrong that most philosophers treat intuitions as conferring any positive epistemic status (see e.g. Cappelen 2012; Deutsch 2015b). Note that question is not whether philosophers should treat intuitions this way, but whether they do so. If the philosophers do not, then the arguments from experimental philosophy are irrelevant (Cappelen 2012, 2014c; Deutsch 2015b). If philosophers do treat intuitions this way, however, they need to worry about the criterion of intuitive fit not being a plausible theoretical desideratum and, in turn, possibly worry about their methods. I will argue that political philosophers treat intuitions as conferring some positive epistemic status.

\textsuperscript{8} In Section 3.3 I provide a brief overview of the findings from experimental philosophy studies.

\textsuperscript{9} Looking ahead, this does not exclude the possibility that the empirical premise may be rejected in the light of considerations other than those just mentioned. Thus, I do not grant that the evidence for the premise is conclusive or that the premise is indubitable. To illustrate, although the empirical premise may be fine when considered in the light of the available empirical evidence, other non-empirical considerations (e.g. moral considerations) may count against it. For example, Thomas Nagel (2012, 106) rejects the evolutionary account of moral intuitions in the light of metaethical considerations. Noting that the evolutionary account of moral intuitions and moral realism are incompatible, he rejects the former. This suggest that while Nagel does not question whether empirical evidence support the evolutionary account of moral intuitions, he thinks that the evidence for moral realism is comparatively stronger. According to such a view, evidence may be transferred across domains. This view will be elaborated and discussed in more detail in Part 2 of the thesis.

4
INTRODUCTION

Therefore, the second aim of the thesis will be to investigate whether the criterion of intuitive fit is essential to the current mainstream philosophical method. This amounts to considering whether the mainstream method can be interpreted so that it does not necessarily rely on the criterion of intuitive fit as a theoretical desideratum. It is commonly claimed that the most commonly applied method in political philosophy (Floyd 2017b; List and Valentini 2016, 542) and moral philosophy (de Lazari-Radek and Singer 2014, 98; DePaul 2011, 2013b; Kappel 2006, 131, 133; McMahan 2013, 110; McPherson 2012, 531; 2015, 652; Tersman 2018) more generally, is the method of reflective equilibrium (MRE). Some even claim that it is the only viable method (Copp 2012, 33; DePaul 1993, 1998, 2006; Lycan 2019, 120; Pogge 2007, 176; Scanlon 2002, 149). Since MRE is widely considered the mainstream method within moral and political philosophy, I will henceforth consider the challenge from experimental philosophy as a challenge to MRE (see e.g. Paulo 2019, 2020; Singer 2005). Very roughly, according to a common explication of MRE, we should build theories by bringing intuitions at all levels of generality, principles, and background theories into coherence (Daniels 2020).

The second aim of the thesis, then, is to investigate whether the criterion of intuitive fit plays an essential role in MRE. Even if it should turn out that MRE, according to the mainstream interpretation of the method, treats the criterion of intuitive fit as a theoretical desideratum, it does not necessarily follow that MRE on any plausible interpretation so treats the criterion of intuitive fit. However, if the criterion of intuitive fit is essential to MRE, then any interpretation of MRE must treat it as a theoretical desideratum. If it is the case that the criterion of intuitive fit plays an essential role in MRE, then the challenge from experimental philosophy is successful: There is no (valid) interpretation of MRE according to which the criterion of intuitive fit does not play a role. Thus, MRE must be rejected.

It is worth stressing some points concerning how the two aims of the thesis relate to each other. It might seem that if it is true that (i) political philosophers treat intuitions as conferring positive epistemic status, and it is true that (ii) most of them apply MRE, then the answer to the second question of the thesis seems to be already given. The criterion of intuitive fit seems to be essential to MRE. However, this need not necessarily be the case. It might be the case that MRE is compatible with the criterion of intuitive fit, while it is also true that the criterion of intuitive fit is not an essential part of MRE. Consequently, if MRE is merely compatible with the criterion of intuitive fit, it could be the case that most philosophers applying MRE apply it as if the method presupposes the criterion of intuitive fit, although—since MRE does not

---

10 In fact, it has been claimed that MRE is the method applied by philosophers in general (DePaul 1998, 294).
actually presuppose it—MRE could be applied without relying on the criterion of intuitive fit. In other words, if MRE is merely compatible with the criterion of intuitive fit, it is possible to reject the criterion of intuitive fit without rejecting MRE.

One might wonder why what philosophers do when they apply method may involve doing something which is not required by the method at a more fundamental level. In the case of MRE, it is worth noting that the method often is considered to be an unattainable ideal (see e.g. Brink 1989, 131, 142; DePaul 1993, 22; 2011, xcii, ciii; Knight 2017, 50; Sayre-McCord 1996, 142; Scanlon 2002, 141; Tersman 1993, 45). According to this assertion, the proponents of MRE will have to settle with a practical approach that does not fully realize the ideal when they seek to apply the method in practice. Thus, when applying MRE, one may accept the criterion of intuitive fit out of convenience.

This suggests that we can distinguish between two ways of understanding the challenge from experimental philosophy: it can be understood as (1) a challenge to the ideal of MRE or (2) a challenge to the way MRE is (currently) applied. According to (2), then, even if there are significant problems suggesting that contemporary philosophers cannot apply MRE the way they currently apply it, there could still be ways of applying MRE that do not suffer these problems.

To summarize, granting to some extent the argument from experimental philosophy, this thesis poses the following questions:

Overarching question: Does the challenge from experimental philosophy succeed?

I. Do contemporary analytic political philosophers treat the criterion of intuitive fit as a theoretical desideratum?

II. Does MRE presuppose the criterion of intuitive fit?

(I) and (II) relate to the overarching question in the following way. If (I) is answered in the negative, then, assuming that political philosophers apply (a valid interpretation of) MRE, the challenge from experimental philosophy fails. If political philosophers apply MRE, but do not treat the criterion of intuitive fit as a theoretical desideratum, then the arguments from experimental philosophy do not seem to have any bearing on the plausibility of MRE. However, if (I) is answered in the affirmative, then we must proceed to investigate (II). I argue that
political philosophers treat the criterion of intuitive fit as a theoretical desideratum. Thus, scrutiny of (II) is warranted.  

If (II) is answered in the affirmative, the challenge succeeds. MRE must be rejected. Answering (II) in the negative, amounts to showing that although contemporary analytical philosophers treat the criterion of intuitive fit as a theoretical desideratum, this criterion is not an essential part of MRE. This would entail that there are implementations of MRE that do not rely on the criterion of intuitive fit as a theoretical desideratum. In other words, while answering (II) in the negative would not mean that the challenge from experimental philosophy does not undermine the current implementation of MRE, it would mean that the challenge fails to undermine the ideal of MRE. While suggesting there is reason to believe that challenge undermine the mainstream implementation of MRE, I argue that the criterion of intuitive fit is not an essential feature of (some interpretations of) the ideal of MRE. Thus, political philosophers need not reject MRE altogether, but they have to (given the assumptions I have granted) revise how they implement the method. They must derive an implementation from the ideal according to which the criterion of intuitive fit is not treated as essential.

These remarks highlight why it is interesting to investigate (I) and (II) within the same thesis. By investigating (I) and (II), the thesis demonstrates that what proponents of MRE do might differ from what the ideal of MRE essentially requires. Moreover, by investigating both (I) and (II), the thesis contributes to explaining why the criterion of intuitive fit might appear essential to MRE, although it is not. This is an important observation. To illustrate, if those who challenge MRE target how MRE is currently applied and those who defend MRE have in mind a more abstract conception of MRE when responding to their critics, then the parties would be talking about slightly different things.

This thesis consists of two parts: an introductory part and a part consisting of three articles. The introductory part contains reviews of the relevant strands of literature and explains how the thesis draws on and contributes to those literatures. Moreover, it also includes discussions of questions from the philosophy of science and research ethics that are relevant to the

11 It could be argued that (II) would be worth investigating even if (I) is answered in the negative. We might want to investigate whether philosophers committed to MRE actually philosophize according to MRE’s prescriptions. If philosophers fail to follow the prescriptions of MRE or even, contrary to what is commonly believed, apply a completely different method, then it is still an open question whether MRE presupposes the criterion of fit, even if (I) is answered in the negative. However, since I assume that the method most philosophers apply is an implementation of MRE and answer (I) in the affirmative, the possibility presented in this footnote will not be considered in this thesis.
IN\-TR\-OD\-U\-C\-TION

despite the investigation of (I) and (II). Finally, it offers brief presentations of the articles collected in the second part of the thesis. The second part of the thesis contains the three articles.
PART 1: BACKGROUND AND GROUNDWORK FOR THE THESIS
Overview of Part 1

Part 1 of the thesis aims to provide a background for the research questions and explain how this thesis contributes to the literature on which it builds. As explained in the introduction, there are two ways of understanding the challenge from experimental philosophy: (1) it can be understood as targeting the ideal of MRE or (2) as targeting how MRE is currently applied. To clarify the purpose of the current thesis and to explain how the two ways of understanding the challenge relate to each other, it is helpful to consider the following claims:

A. Most political philosophers apply MRE
B. MRE can only be a valid method if intuitions can be treated as data
C. According to the findings of experimental philosophy studies, intuitions are not suited for serving as data for moral theorizing
D. Most political philosophers treat intuitions as conferring some positive epistemic status

As mentioned in the introduction, A is a common assertion. B is an elaboration of the assertion that MRE presupposes the criterion of intuitive fit. It says something about why intuitive fit is a theoretical desideratum. By the claim that intuitions are supposed to serve as data, I mean that intuitions are supposed to ensure that our theories represent what we want them to represent. For example, if we want our theories to represent the facts about a particular matter, intuitions are supposed to be a source of information about those facts.

It is worth noting that B, in contrast to A and D, is not a claim about what philosophers do, but a theoretical claim about MRE. B asserts that the criterion of intuitive fit is essential to MRE. If intuitions cannot be treated as data, as C asserts, there is no point in trying to make our theories fit our intuitions. Since applying MRE (supposedly) entail (at least to some extent) conforming our theories to our intuitions, it would be a severe blow to the method if it should
turn out that intuitions cannot be treated as data. If the method relies on intuitions as the only, or a necessary, source of making sure that theories represent what we want them to represent, then it would follow from C that the method is flawed.

C is an explication of the arguments from experimental philosophy. To illustrate C, if the assumption is that we want our theories to represent the facts about a particular matter, the arguments from experimental philosophy amount to the objection that our intuitions do not (reliably) track those facts.

Finally, D is also a common assertion. I will assume that philosophers treat intuitions as conferring some positive epistemic status because they assume that intuitions track what we want our theories to represent (e.g. moral facts)—i.e. most philosophers treat intuitions as data for moral theorizing.

The first formulation of the challenge from experimental philosophy presupposes C and B. If B is true, then either A or D must false.\(^{12}\) The second formulation presupposes A, D, and C. From A and D it follows that according to current implementations of MRE, intuitions are treated as data for moral theorizing. Since the challenge target current implementation of MRE, B is not that relevant. It does not matter that there might be an implementation of MRE according to which intuitions are not treated as data.

If the first formulation succeeds, then the challenge against the current implementation succeeds as well. Any valid implementation of MRE must treat intuitions as data. If D should turn out to be false, this would mean that most political philosophers do not apply (a valid interpretation of) MRE. That is, if B is true it follows that philosophers who apply (a valid interpretation of) MRE treat intuitions as data. However, the opposite does not hold. If the second formulation of the challenge succeeds, it does not necessarily follow that the first formulation succeeds. Although the current implementation of MRE involves treating intuitions as data, this does not mean that the ideal of MRE necessarily presupposes that intuitions are suited for serving as data. It could be possible to derive a practical approach from the ideal of MRE according to which intuitions are treated as data even if treating intuitions as data is not an essential feature of MRE according to the ideal. That is, B might still be false.

The main motivation for asking the overarching research question is to investigate whether MRE can withstand the challenge (presumptively) posed by the findings from experimental philosophy studies. Of the claims listed above, A will simply be asserted. As explained in the introduction, C will, to some extent, be granted. That is, the empirical basis of the argument

\(^{12}\) Or they apply an invalid interpretation of MRE. However, since I argue that D is true, the current thesis does not consider this possibility. (See also footnote 11 in the introduction.)
from experimental philosophy will not be questioned. I will assume that when we consider the empirical premise in light of the available empirical evidence, the premise is fine. Moreover, I will, for part of my investigation, grant the argument from experimental philosophy for the sake of argument.\textsuperscript{13}

This leaves us with mainly B and D to investigate, hence question (I)—Do contemporary analytic philosophers treat the criterion of intuitive fit as a theoretical desideratum?—and (II)—Does MRE presuppose the criterion of intuitive fit? As the overview has demonstrated, answering (I) is crucial for assessing whether the version of the challenge that targets the current implementation of MRE succeeds and answering (II) is crucial for determining whether the challenge targeting the ideal of MRE succeeds.

An important aim of Part 1 of this thesis is to review the strands of literature related to these claims and the challenge from experimental philosophy more generally and explain how this thesis contributes to that literature. In Section 2, I review the literature on moral intuitions and the method of reflective equilibrium. In Section 3, I review the literature on experimental philosophy and outline the challenge from experimental philosophy in more detail. In Section 4, I review the literature on the debate concerning whether intuitions are treated as conferring positive epistemic status, or data, in philosophical inquiry. Having reviewed the relevant strands of literature, Section 5 explains how this thesis contributes to the ongoing debates. Moreover, in Section 6, I discuss the relationship of this thesis to the philosophy of science and some relevant issues concerning good research ethics. Finally, I offer a brief presentation of the articles in Section 7.

\textsuperscript{13} I respond to worries about making these assumptions regarding C in Section 3.7.
Moral Intuitions and the Method in Moral Inquiry

In this section, I review the literature on moral intuitions and claims commonly made about MRE. Since MRE is the target of the challenge from experimental philosophy, and the challenge relies on the assumption that intuitions play an essential role in MRE, this is a natural place to start this literature review.

2.1 Moral Intuitions

I start this section by outlining what intuitions are and some features commonly attributed to intuitions. I then distinguish between different types of intuitions and account for what role moral intuitions are assumed to play in moral inquiry.

2.1.1 What are Intuitions?

Intuition is a certain sort of mental state (Stratton-Lake 2016, 28). There are different views on what the nature of this state is. Three of the more common accounts are the following: intuition is a form of belief, intuition is an inclination to believe, and intuition is an intellectual seeming (Stratton-Lake 2016; see also Pust 2016, ch. 2; 2019).

1 This list of accounts is not exhaustive. For example, John Bengson (2015) argues that intuition is a form of quasi-perception.

The case commonly referred to as “Footbridge” may help illustrate these views. In Footbridge, a runaway trolley is about to run over and kill five people. A bystander standing on a footbridge can stop the trolley from running over the five people only by pushing a very large stranger, who is also standing on the footbridge, onto the track, bringing the trolley to a halt. However, pushing the stranger in front of the trolley will kill the stranger. Is it permissible for the bystander to push the stranger in front of the trolley? (Thomson 1985, 1409).
On the first view, your having the intuition that it is impermissible for the bystander to push the very large stranger, means that you believe that it is impermissible to push the large stranger. On the second view, your having the intuition that it is impermissible to push the large stranger means that you are inclined to believe that it is impermissible to push the large stranger. On the third view, if you have the intuition that it is impermissible to push the stranger, then it seems—or appears—to you that it is impermissible to push the large stranger.

To see that it may seem to a person that is impermissible to push the stranger, but that the person need not believe or being inclined to believe that it is, consider a committed act utilitarian who believes that we ought to do what maximizes the overall utility. Although the utilitarian does not believe, or is even inclined to believe, that it is impermissible to push the stranger, it still might seem to the utilitarian that it is impermissible to push the large stranger.

According to Matthew S. Bedke (2019, 362), there is a growing consensus that the latter account—that intuition is an intellectual seeming—is the correct account (see also McPherson 2020, 44). The arguments of this thesis do not depend on a particular understanding of the nature of intuitions. I take it that the accounts of the nature of intuitions discussed above are attempts at accounting for the same phenomenon. What is relevant to the current inquiry is what the role of the phenomenon is, not what its nature is (see also Section 4.1). However, for the ease of exposition, and because it is the most popular account, I will assume that intuitions are intellectual seemings in the remainder of the thesis.

A feature commonly emphasized in accounts of intuitions, and commonly accepted among proponents across the aforementioned accounts, is that intuition is formed non-inferentially (Jenkins 2014, 94; Liao 2016, 319; Osbeck 1999, 234; Osbeck and Held 2016, 6-7). That is, the seeming to one that [e.g. it is not permissible to push the stranger in footbridge] is just such as the belief that it is not permissible to kill another human being. It just seems to you that it is impermissible to push the stranger. Following C.S.I. Jenkins (2014, 94), we may say that intuitions are “psychologically non-inferential.”

2.1.2 The Epistemic Status of Intuitions

Another quite common assertion is that intuitions justify non-inferentially (Bedke 2008; 2010, 1069; Jenkins 2014, 94). If you have the intuition that it is impermissible to push the stranger off the footbridge, then you have some degree of justification in believing that it is impermissible to push the large stranger.
impermissible to push the stranger off the footbridge. This justification is not derived from any moral principle and it is independent of whether the claim that it is impermissible to push the large stranger off the footbridge can be supported by any such principle. Following Jenkins, we may say that intuitions are “epistemically non-inferential” (Jenkins 2014, 94).³

The distinction between “intuitings” and “intuiteds” (Lycan 1986, 88-89; 1988; see also Lycan 1985, 114) helps clarify what we mean when we say that intuitions justify beliefs. “Intuiting” refers to the occurrence of the mental state. “Intuited” refers to the content of the mental state. To illustrate, if I have the intuition that it is impermissible to push the very large stranger, the occurrence of the intuition is the intuiting, whereas the content of the intuition—that is, the proposition that it is impermissible to push the large stranger—is the intuited. When we say that intuitions confer justification onto beliefs, we mean that “the occurrence of an intuiting is taken to provide the person in whom it occurs with prima facie justification for believing the intuited” (Pust 2019, italics in original; see also Tersman 2012, 644). In terms of our example, that it seems to you that it is impermissible to push the stranger provides you with some justification for believing that it is impermissible to push the stranger.

The assertion that intuitions provide non-inferential justification in this way is not uncontroversial (see also Lemos 2019, 384-385). It is common to separate between two major accounts of justification: foundationalism and coherentism. According to foundationalism, a belief is justified if it is non-inferentially justified or stands in the appropriate relationship with a non-inferentially justified belief (Lammenranta 2004; Lemos 2019; McMahan 2013, 110). This account of justification, then, is compatible with the assertion that intuitions are epistemically non-inferential. Intuitions, on this account, can be one source of non-inferential justified beliefs. According to coherentism, on the other hand, whether a belief is justified depends entirely on its relation to other beliefs (Lammenranta 2004; Lemos 2019; McMahan 2013, 110). Thus, beliefs cannot be non-inferentially justified. According to this view, beliefs based on intuitions are not non-inferentially justified; their justification depends on their coherence with other beliefs—for example, our second-order beliefs about our intuitions (Brink 1989, 117-136; Daniels 1996, 83; Sayre-McCord 1996, 160-161; Tersman 2008, 399; 2018, 4). For example, beliefs based on intuitions may be justified because believing that beliefs based

³ For an overview of different epistemological theories of how intuitions non-inferentially justify beliefs, see (Bedke 2010).
on intuitions are formed reliably increases the coherence of our system of beliefs. Thus, if intuitions confer justification, they do so inferentially.

To summarize, the assertion that intuitions confer non-inferential justification is common but not universally accepted. Coherentists, as we have seen, do not think that intuitions justify non-inferentially. On the other hand, coherentism is compatible with the assertion that intuitions are psychologically non-inferential.

2.1.3 Different Types of Intuition

We may distinguish between different types of intuitions based on the level of generality of their content (see e.g. Greene 2014, 724; Huemer 2008; Kagan 2001; McMahan 2013, 104-105; Sandberg and Juth 2010; Singer 2005; Sinnott-Armstrong, Young, and Cushman 2010, 247). There are different ways of drawing the distinctions between different types of intuitions based on their generality. The number and names of the categories vary between the accounts. However, these accounts all represent the same underlying idea. In this thesis, I will follow Michael Huemer (2008, 383) and divide intuitions into three groups based on their level of generality:

1. **Concrete intuitions:** These are intuitions about specific situations, such as the intuition that in Singer’s Shallow Pond example, one is obligated to rescue the drowning child, or that in [Footbridge, one is not permitted to push the very large stranger…].

2. **Abstract theoretical intuitions:** These are intuitions about very general principles, such as the intuition that the right action is always the action that has the best overall consequences, or that it is wrong to treat individuals as mere means.

3. **Mid-level intuitions:** These are intuitions about principles of an intermediate degree of generality, such as the principle that, other things being equal, one ought to keep one’s promises; that one ought to show gratitude for favors done to one; or that it is more important to avoid harming others than it is to positively help others. (Italics in original, note removed)

Shelly Kagan (2001, 61; see also Bennett 1993, 471) suggests that distinguishing between different types of intuitions based on the different kinds of objects for intuition might be somewhat misguided:

when we react to particular cases we are actually reacting to things of the very same type as when we react to general moral claims. It is easy to lose sight of this, given our common practice […] of saying that we are reacting to particular cases. But what we are actually reacting to, I think, are types of cases.
This is easiest to see in the situation where the kind of case we are thinking about is purely imaginary. What we are presented with, then, is only a description—and typically, all things considered, a fairly thin description at that. There is no actual, particular, concrete case that we are confronted with. So when our intuition tells us, say, that some particular act would be the right thing to do in that particular case, what we are actually intuiting, it seems, is that a certain kind of act would be the right thing to do in a certain kind of case. And this, of course, is a general moral claim. (Italics in original)

Even if we grant Kagan’s claim that we never consider concrete cases but rather thin descriptions of cases, I think there remains an important difference between judging thin descriptions of cases and judging principles. I do not think that Kagan’s point amounts to an objection to Huemer’s taxonomy. It might call for a revision of the label used to refer to the first category of intuitions (and the addition of a fourth category (see McPherson 2020, 39)). Still, as long as considering cases is an activity distinct from considering whether a principle has intuitive appeal on its own, I think it makes sense to single out intuitions about cases as a distinct type of intuitions.4

2.2 The Role of Intuition in Moral Inquiry

It is commonly claimed that in moral philosophy, intuitions are used to construct and test moral theories (Bedke 2019, 361, 363; Shaw 1980; see also Dworkin 1989, 27-29; Elster 2011, 242-243; Kagan 2001, 45).5 Roughly, the idea is that if I have the intuition that p, I should when constructing a moral theory, try to construct a moral theory that is consistent with p—i.e., the content of my intuition (Bedke 2019, 363; Tännö 2015, 10; see also Cath 2012). Moreover, I can test a moral theory by considering how it fits my intuitions—i.e., the content of my intuitions. That a theory is consistent with my intuitions is something that counts in favor of the theory. That my intuitions conflict with the theory is something that counts against the theory. Of course, it is seldom the case that all our intuitions are either consistent or conflict with a theory. Thus, the more of our intuitions that are consistent with a theory, the stronger reason we have to believe that the theory is implausible. Conversely, the more of our intuitions that conflict with a theory, the stronger reason we have to believe that the theory is implausible.

4 By the claim that contemplating cases is a distinct activity from considering whether a principle is intuitively appealing, I mean that when contemplating cases, we do not typically judge the cases by identifying what principles the different options in the case imply and base our judgment about the case on which principle we find (the most) intuitively appealing. I take this assertion to be rather uncontentious and compatible with Kagan’s claim. Looking ahead, the upshot of this distinction is that objections that apply to one type of intuition need not apply to another type of intuition (see e.g. Huemer 2008).

5 For a defense of a related claim about philosophy more generally see (Pust 2016, ch. 1).
Intuitions are often asserted to play a role in moral philosophy analogous to the role observations play in science (Kagan 2001, 127, 130; McMahan 2013, 114; Shaw 1980; Tännsjö 2007). An assumption that motivates this picture of the role of intuitions, is that intuitions are theory-independent (Audi 1997, 41; Huemer 2005, 103; Kamm 2001, 8-9; Kauppinen 2018, 52; Tännsjö 2015, ch. 1; see also Andow 2019, 266; Goldman and Pust 1998, 183; Hitchcock 2012, 207; Jenkins 2014, 92). This assumption is related to the assumption that intuitions are psychologically (and epistemically) non-inferential. We do not form intuitions by inferring claims from theories. This kind of independence make intuitions suited as data for testing theories.\(^6\)

A further assumption that underlies the picture that the role of intuitions in moral theorizing is analogous to the role of observation in science is that intuitions track what we want our moral theories to represent. That is, when we construct a moral theory “there is something the theory is about: any theory is intended to represent, summarize, or capture something ‘outside the theory’” (List and Valentini 2016, 540). We might say that a theory is externally valid if it represents what we want it to represent (List and Valentini 2016, 538).\(^7\) Intuitions, then, are suited as data for moral theorizing, given the assertion that they track what we want our theories to represent.

It might not be that clear what we want our moral theories to represent, and one might wonder whether the purported analogy holds across different metaethical positions. As Christian List and Laura Valentini (2016, 540-541) notes:

\[
\text{If we are realists about normative or evaluative matters, we may say that the theory is intended to represent some theory-independent moral facts. If we are not realists about normative or evaluative matters, it is harder to specify what a normative or evaluative theory is intended to represent. [...] Yet the very idea of a theory breaks down unless we assume that there is something potentially representable by it, however observer-dependent or socially constructed it might be.}
\]

---

\(^6\) To say that intuitions are completely theory-independent might be too strong. Some claim that our intuitions, like our observations, are theory-laden (see e.g. Føllesdal 1979, 331-332). However, this still allows for some kind or degree of theory-independence. That is, it does not necessarily follow that the content of intuitions is solely derived from the theories that they are supposed to test (see also Bengson, Cuneo, and Shafer-Landau 2022, 39-41).

\(^7\) Christian List and Laura Valentini (2016, 538-544) distinguish between two kinds of criteria that we may use to assess moral principles and theories: internal and external criteria. Internal criteria “concern the way the principles or theory are formulated and their internal logical structure,” while external criteria “concern the relationship between the principles or theory and what these are ‘about’: their normative or evaluative content” (List and Valentini 2016, 538).
Thus, if we are moral realists, we want our moral theories to represent moral facts. On this view, intuitions serve as data for moral theorizing on the assumption that they track moral facts (see e.g. Huemer 2005; Shafer-Landau 2003; see also Dworkin 1989, 27-28).

Moreover, although it is not that clear what we want our theories to represent if we are not realists who believe there are theory- or mind-independent moral facts, and thus whether the analogy holds in that case, the fact that philosophers who are not realists in this sense also engage in theory building, seems, at least according to List and Valentini, to suggest there is something representable by their moral theories. Moreover, it seems commonplace to assume that intuitions track this “something.” For example, according to moral constructivism, insofar as there are moral truths, they are not fixed by (mind-independent) moral facts (as realists think), but constituted by what rational agents would agree to under specific circumstances (Bagnoli 2021). On this view, the assertion is that our deep convictions determine what rational agents would agree to under the specific circumstances, and our intuitions reflect our deep convictions (Ebertz 1993, 208; see also Dworkin 1989, 28-29).

Moreover, a nonrealist about morality may consider the aim of moral philosophy to be to construct a moral theory that organizes one’s commitments. Pursuing such an aim could be viewed as “personal choice” reflecting one’s “wish to be whole and interconnected in [one’s] person, so that [one] can understand some of [one’s] attitudes as parts or upshots of others” (Bennett 1993, 471; see also Raz 1982, 310-311). In that case, intuitions are treated as data because they represent one’s commitments.

To summarize, when constructing moral theories, there is, regardless of our metaethical views, something we want our theories to represent. Most metaethical views are compatible with the view that intuitions track that something that we want our theories to represent and therefore can be used as a means to construct and test theories (Sayre-McCord 2013; see also Bengson, Cuneo, and Shafer-Landau 2019; Copp 2012).  

2.3 The Method of Reflective Equilibrium

In this section, I outline what I take to be common claims about MRE. Moreover, I also indicate some aspects of the method about which its proponents have different views. As mentioned in the introduction, MRE is claimed to be the most widely applied method for doing political

---

Matthew S. Bedke (2009, 205) suggests that even if it should turn out that some metaethical positions are incompatible with intuitive justification (as he thinks is the case), we should reject those metaethical positions rather than intuitive modes of justification: “For my own part, I think intuitive modes of justification are indispensable in ethics, and so we should explore other metaphysical views that leave the intuitive justification intact.”
philosophy and moral philosophy more generally. The general idea behind the method is that one revises one’s relevant input until one achieves a coherent set of beliefs. (I will shortly return to what counts as “relevant input” according to MRE.) I will refer to the end state of MRE—that is, roughly, the state one achieves when one’s beliefs form a coherent whole—as “reflective equilibrium” (RE). Following Roger P. Ebertz (1993), I will refer to the process by which one reaches RE as “the reflective process.”

2.3.1 Narrow and Wide MRE

It is common to distinguish between a narrow and a wide version of MRE (see e.g. Cath 2016; Daniels 1979, 2020; Rawls 2001, 30-31). How the distinction is drawn varies somewhat. However, the important point is that the wide version of MRE is more inclusive in the sense that it regards a wider variety of beliefs as relevant to moral inquiry. According to the narrow(est) construal of MRE, one treats one’s considered moral judgments (CMJ) and the principles that do not deviate to much from these judgments as input to the method (Rawls 1999, 43; see also DePaul 1986, 59; Knight 2017, 49). If conflict among these beliefs is discovered, one revises either CMJs or principles to resolve it. When all conflicts are resolved, and one’s set of beliefs is coherent, one has reached a narrow RE.

Wide MRE, which is the version of MRE that most proponents of MRE embrace (Cath 2016, 218; Rawls 1999, 43; 2001, 31; but see Holmgren 1989), is wider than narrow MRE in several respects. First, the moral principles included in the reflective process are not restricted to only moral principles that do not deviate too much from one’s CMJs. Second, wide MRE also includes background theories as input to the reflective process. The term “background theories” is often used quite loosely (DePaul 2011, lxxxii; see e.g. Daniels 1979, 258) and a wide variety of theories is suggested as examples of what may be relevant background theories. “Background theories” need not be full-fledged theories but can be more or less generic claims concerning some area of interest (DePaul 2011, lxxxii). Thus, there is no ready-made simple definition of background theories. However, in the following I will account for each element in some more detail and eventually clarify the concept of background theories. Since it is the wide version most proponents of MRE embrace, it is the wide version that is the primary concern of this thesis. Thus, unless otherwise noted, when I use MRE in the remainder of the thesis, I have the wide version of MRE in mind.

It might be more accurate to say wide versions of MRE. I return to this point below.
2.3.2 Considered Moral Judgments and Intuitions

CMJs are moral judgments at all levels of generality (Rawls 1974, 8; 2001, 30) that satisfy certain criteria (Rawls 1999, 42; see also Daniels 1979, 258; DePaul 1993, 16-17). Commonly, these criteria are assumed to amount to the judgments being made under circumstances favorable for making judgments in general. For example, Rawls (1999, 42; 2001, 29) maintains that CMJs are judgments that are not given when upset or frightened, that we have some confidence in, and that are made by a person with the ability and desire to reach a correct decision. However, according to some proponents of the method, the criteria used to decide which judgments qualify as CMJ may themselves be revised in the reflective process (DePaul 1993, 17; see also Tersman 2008, 398).

It is worth briefly commenting on how intuitions are related to CMJs. Often, the terms “CMJ” and “intuition” seem to be used interchangeably. That is, some philosophers write that intuitions, rather than CMJs, are treated as input to the method (see e.g. Appiah 2008, 75; McMahan 2013). This could either suggest that “CMJ” and “intuition” are used synonymously in the literature on MRE or that CMJs and intuitions are not the same things and that there is some disagreement regarding what the proper input to the method is. In this thesis, I will assume that “CMJ” and “intuition” are used more or less synonymously and designate the kind of mental state described in Section 2.1. Thus, henceforth I will not use “CMJ” but “intuition” to refer to this element of the method. By doing so, I bracket one potential response to the challenge of experimental philosophy; the proper input of MRE is not intuitions, but CMJs (see e.g. Brun 2014; Greenspan 2015, 108-109; Sandberg and Juth 2010, 221-222; Tersman 2008, 391-392, 397-398; see also Copp 2012, 18-19).\(^{10}\) If this assertion is wrong—that is, if “CMJ” and “intuition” are not referring to the same thing—I do not think that would leave this thesis irrelevant, although I assume that intuitions (as accounted for in Section 2.1) are (part of) the input of MRE. A considerable number of philosophers assert that intuitions are what is treated as input according to MRE (see e.g. Audi 1997, 50 n. 37; Brink 2014, 687; Huemer 2005, 117; Pust 2016, 13; Singer 2005).

2.3.3 Moral Principles

According to Christian List and Laura Valentini (2016, 535),

\(^{10}\) Jonathan Floyd (2017a, 123) makes a distinction between moral intuitions and CMJs, although in a slightly different context. Floyd is not making this distinction to defend MRE from experimental arguments.
Moral principles are principles which content is moral. Moral principles can be more or less general. As noted in the discussion of intuitions, we have intuitions at all levels of generality. Thus, some principles are supported by intuitions. However, principles without intuitive support may also play an important role in the reflective process. For example, although a principle might lack intuitive appeal on its own, it could be grounded in an even more general principle (or rationale) or supported by background theories (Daniels 1979, 258-259). If a principle is well-grounded in this way, we may be inclined to revise intuitions about cases when these conflict with the principle, despite the principle’s lack of intuitive appeal. Moreover, although a principle might lack intuitive support itself, it might gain credence in the reflective process due to its ability to systemize and account for our concrete intuitions. It could be the case that accepting a non-intuitive principle could increase the coherence of our system of beliefs.

2.3.4 Background Theories

The term “background theory” is often used quite loosely and encompasses theories (or assertions) from a wide variety of domains. These domains include domains such as “metaphysics (e.g., views concerning free will, or the nature of persons or actions), epistemology (e.g., views about the kinds of circumstances where we’re likely to make mistakes or the possibility of a priori justification), language (especially the analysis of moral language or accounts of the etiology of moral terms), and also elements of psychology, sociology (e.g., regarding the role morality plays in society), and biology” (DePaul 2013b).

According to Daniels (1979, 258), background theories provide us with resources for bringing out the relative strength and weaknesses of competing moral principles. That is, we can construct arguments for or against moral principles as inferences from some set of relevant background theories. The following example provided by Folke Tersman (1993) may be illustrative. Tersman (1993, 57), drawing on Parfit (1984), claims that views about personal identity may serve as background theories that may have implications for what we should believe about certain moral issues. For example, if one’s view of personal identity implies that “we are right to ignore whether experiences come within the same or different lives” (Parfit 1984, 206 cited in Tersman 1993, 58), that view supports utilitarianism, since utilitarianism...
calls for maximization of welfare with no regard to whether the welfare is experienced within the same or different lives.

It is worth noting that proponents of MRE seem to differ somewhat with regard to which background theories they think should enter the reflective process. For example, while Rawls (1974) (initially) wanted to distance moral theory from metaethics, others have argued that metaethical considerations bear on first-order moral questions in a way that makes our inquiry suffer if such considerations are not included (McPherson 2012).

2.3.5 MRE and the Reflective Process

I have now described the different elements that are treated as input to MRE: moral intuitions, moral principles, and background theories. A proponent of MRE is supposed to build a moral theory by bringing these elements into coherence. As I have already mentioned, this entails that conflicts among these elements must be resolved by revising the elements under considerations until there is no longer any conflict among them. In the reflective process, no element is immune to revision (Rawls 1974, 8; 1999, 42; 2001, 30-31). According to MRE, one should resolve such conflicts by making the revisions that increases the coherence of one’s set of beliefs. However, sometimes considerations of coherence will not help us decide what revisions to make (see e.g. Haslett 1987, 310). The alternative sets of beliefs under consideration could be equally coherent. In that case, the method recommends that the person applying MRE make the revision that on balance seems most true to her (DePaul 2011, lxxv-lxxxi; 2013b).

We may, then, describe MRE as a method consisting of the following steps (see also Cath 2016; Daniels 1979, 258-259): First, we identify a set of relevant intuitions about cases. Second, we identify relevant (intuitive and non-intuitive) moral principles and background theories. In the third stage, we move back and forth between the three sets of beliefs, trying to bring them into coherence. Each set of beliefs can be revised to resolve conflict within the sets and to fit the other sets of beliefs. Note that we may, in the reflective process, move back and forth between the different stages. For example, we may introduce further intuitions about cases later in the process because we were unaware of them when initially engaging in the reflective process or we initially failed to recognize them as relevant.

With this description of MRE in mind, it is worth noting that MRE is often said to be analogous to the method in science (Cummins 1998, 113; Sayre-McCord 1996, 141-142; Shaw 1980, 130). Thus, while the role of intuitions in moral inquiry is analogous to the role of observations in science, MRE is analogous to the procedure by which one systematizes
observations and constructs theories based on them in science (see also Føllesdal 1979, for a relevant discussion).

Another point worth noting, and which is of particular importance for the purpose of the current thesis, is that MRE is quite demanding. In fact, many of its proponents consider MRE to be an ideal that is not humanly attainable (see e.g. Brink 1989, 131, 142; DePaul 1993, 22; 2011, xcii, ciii; Knight 2017, 50; Sayre-McCord 1996, 142; Scanlon 2002, 141; Tersman 1993, 45). Thus, when proponents of the method seek to apply the method in practice, they will have to settle with a practical approach that does not fully realize the ideal.

### 2.3.6 Some Variations

Although I have tried to characterize MRE in a way that should be rather uncontroversial, there is some variation regarding how proponents of MRE interprets the method. I have already hinted at one way in which interpretations may differ. That is, interpretations may vary with regard to which domains are considered relevant to collect background theories from. Moreover, different interpretation may vary with regard to how extensive the set of background theories considered in the reflective process is.

Another point of disagreement is whether MRE is committed to epistemic coherentism, epistemic foundationalism, or any theory of epistemic justification at all (see e.g. DePaul 1986, 2013a; Ebertz 1993; Tersman 2008). Yet another point of disagreement is whether MRE is practiced by individuals (see e.g. DePaul 1993, 22-23; 2011, lxxxi) or is a kind of social endeavor (see e.g. Walden 2013; see also Director 2020). There is also disagreement about whether MRE should make requirements pertaining to the inquirer(s). In earlier writings, Rawls (1951) required that the judgments that served as input to the method should be the judgments of competent judges. However, in later writings (see e.g. Rawls 1999), he does not mention this requirement. Many proponents of MRE seem to follow the later Rawls. However, DePaul (1993, 2006) finds the requirement concerning the inquirer plausible and thinks that part of applying MRE is to pursue formative experiences to educate oneself.

Due to such disagreements, there are (slightly) different interpretations of MRE in the literature. Looking ahead, the differences regarding background theories will be important to the investigation of research question (II).

### 2.4 Taking Stock

Having outlined the central features of MRE and the role intuitions are supposed to play in moral inquiry, we may say something about the role the criterion of intuitive fit—i.e. the
restrictive version of it pertaining to concrete intuitions—(supposedly) play in MRE. The criterion of intuitive fit is treated as a theoretical desideratum because intuitions are assumed to be a source of information about what we want our theories to represent. That is, concrete intuitions are treated as data (see also Bengson, Cuneo, and Shafer-Landau 2022, 36).
In this section, I outline the challenge from experimental philosophy in some more detail. I start by giving a brief outline of experimental philosophy and arguments from experimental philosophy. I then account for the move from arguments from experimental philosophy to the challenge from experimental philosophy.

3.1 Experimental Philosophy

Experimental (moral) philosophy “is the empirical study of moral intuitions, judgments, and behaviour” (Alfano, Loeb, and Plakias 2018; see also Mallon 2016, 410). Philosophers working within this branch of philosophy apply methods from social sciences, in particular, psychology, to study these matters. For the purpose of this thesis, the studies of interest are those on intuitions and how the findings of these studies are used to challenge MRE.

It is worth reiterating that so far, experimental philosophy studies have primarily been concerned with concrete intuitions. Thus, arguments from experimental philosophy challenge mainly a restricted version of the criterion of intuitive fit according to which the intuitions a theory should conform to are concrete intuitions. Therefore, it is this restricted version of the criterion of intuitive fit that is relevant for this thesis.

3.2 Arguments from Experimental Philosophy: The Argument Scheme

Commonly, it is suggested that arguments from experimental philosophy generally take the following form (see e.g. Klenk 2019, 10; Kumar and May 2019, 31; Mukerji 2014, 228; 2018, Ch. 2; Nichols 2014, 731; see also Bengson 2013, for a similar but more detailed scheme):

- Empirical premise: Concrete intuitions have property P
- Epistemic premise: If intuitions have property P, then they cannot be treated as data.
Conclusion: Concrete intuitions cannot be treated as data.

The empirical premise is the premise of the argument that is supported by experimental philosophy studies. In other words, experimental philosophy studies provide us with information about what properties our intuitions have. For ease of exposition, and since the argument suggest that these properties are problematic from an epistemic point of view, I will sometimes simply refer to these properties as problematic properties. Arguments from experimental philosophy can be explicated in numerous ways depending on for example what properties figure in the first premise.¹

The sketched argument scheme serves as a helpful starting point for discussing arguments from experimental philosophy and the challenge from experimental philosophy. In the following, I will elaborate on and expand on this scheme.²

3.3 Arguments from Experimental Philosophy: A Note on the First Premise

In this section, I examine the first premise of the argument from experimental philosophy more closely by outlining some of the properties that intuitions have according to experimental philosophy studies.³

3.3.1 Demographic Variation

The findings of some experimental philosophy studies suggest that people have opposing intuitions about the very same cases and that demographic variables explains this variation. To illustrate, some studies have suggested that we find this kind of disagreement across different cultures (Ahlenius and Tännsjö 2012), gender (Friesdorf, Conway, and Gawronski 2015), age (Hauser et al. 2007), personality (Kahane et al. 2015), political orientation and religious

---

¹ It is not uncommon to distinguish between negative and positive experimental philosophy. According to Ron Mallon (2016, 411),

So-called negative or negative program experimental philosophers view experimental work as revealing that the use of philosophical intuitions as evidence for the truth of a philosophical claim is bankrupt. Other, positive or positive program philosophers see experimental methods as offering the possibility to continue and improve upon traditional uses of intuitions… And still others hope to escape this dialectic altogether… (italics in original)

² For a more detailed overview of such arguments than, and slightly different from, the one I offer in this section, see Hanno Sauer (2018).

³ For overviews of the experimental philosophy studies, see e.g. Mallon (2016); Machery (2017). Note there is some variation with regard to how these overviews, and mine, for that matter, are organized and some studies are categorized under different labels in each of the overviews.
affiliation (Young, Willer, and Keltner 2013), and socio-economic status (Haidt, Koller, and Dias 1993).

3.3.2 Variation Across Situations: Presentation

Some experimental philosophy studies suggest that intuitions about cases are influenced by how the cases are presented. To be more precise, the findings of some studies suggest that intuitions about cases vary based on how the cases are presented, despite no substantive changes to the cases themselves. One example of this kind of manipulation is the manipulation of the order in which the cases are presented (Liao et al. 2012; Schwitzgebel and Cushman 2012, 2015).

To illustrate, in one study (Schwitzgebel and Cushman 2012), the participants were presented with two cases: Switch and Footbridge. In Switch, a trolley will run over and kill five people unless a bystander pulls a switch, in which case the trolley will be diverted onto a different track where it will run over and kill one person instead. In Footbridge, a trolley will run over and kill five people unless a bystander pushes a very large stranger in front of the trolley. The stranger is large enough to stop the trolley; however, if pushed in front of the trolley, the large stranger will be killed by the impact from the trolley. Some participants were first presented with Footbridge, while others were presented with Switch. The findings of these studies suggest that people are more likely to judge the pulling of the leaver in Switch as impermissible when they are presented with Footbridge first, compared to when they are presented with Switch first.

Other manipulations of the presentation of a case that seem to affect our intuitions about it are manipulations of whether the options in a case are framed in terms of saving versus killing (Petrinovich and O’Neill 1996), whether the case is spelled out in third or second person (Nadelhoffer and Feltz 2008), and whether the cases are put in the participants’ native language or a language that is foreign to them (Costa et al. 2014).

3.3.3 Variation Across Situations: Inducing Feelings

Another property of intuitions, according to the findings of some experimental philosophy studies, is being susceptible to the inducement of certain feelings. Experimental philosophers have, for example, looked at whether making participants feel disgust influences their intuitions. They have done so by exposing the participants to bad smells, dirty desks, etc. Some have even tried to induce disgust through hypnosis (Wheatley and Haidt 2005). Participants were hypnotized so that random words that normally do not trigger disgust triggered it. Other studies have investigated whether cleanliness primes—spraying the participants’ surveys with
Lysol—affect moral intuitions (Tobia, Buckwalter, and Stich 2013). Some have investigated the effect of feelings of positivity by showing participants short comedy clips (Valdesolo and DeSteno 2006). Such studies have revealed that the content of our moral intuitions is affected by emotions.

### 3.3.4 Evolutionary History

According to Joshua D. Greene’s dual-process theory of morality, our moral intuitions are the output of affective cognitive processes shaped by our evolutionary history (Greene 2003; Greene 2013, 2014). According to this theory, we have intuitions with a specific content because having intuitions with that specific content enhances our survival. Greene’s theory synthesizes the findings of a large number of experimental philosophy studies.4

### 3.4 Arguments from Experimental Philosophy: A Note on the Second Premise

In Section 2.2, I claimed that intuitions are treated as data in moral theorizing because they are assumed to provide us with information about what we want our moral theories to represent. The assumption underlying the second premise is that having property P suggests that they do not provide us with such information.

Drawing on the distinction between validity and reliability from the literature on measurement (Adcock and Collier 2001), there are two ways of interpreting this assumption. According to the first interpretation, the reason why intuitions should not be treated as data for moral theorizing is that they are invalid. According to the second interpretation, the reason why intuitions should not be treated as data for moral theorizing is that they are (too) unreliable. According to Robert Adcock and David Collier (2001), a measurement procedure is valid when it actually measures what we want it to measure, and it is reliable when repeated applications of it yield consistent results.

Drawing on this distinction, it could be the case that the fact that intuitions have property P is taken to suggest that they lack validity. That is, they do not track what we want them to track. If we are moral realists, we want moral theories that correspond to moral reality. Thus, we want our intuitions to track moral facts. Our intuitions, then, are valid insofar they track moral truth.

To illustrate, the argument from demographic variation—at least in the form of cultural variation—can be formulated as targeting intuitions validity. Cultural variation, it could be

---

4 See also Katia Vavova (2015) for an overview and a discussion of evolutionary debunking arguments.
argued, suggests that our moral intuitions track cultural norms—and such norms might not be what we are after when we engage in moral inquiry. Similarly, evolutionary debunking arguments target the validity of intuitions. The claim underlying this challenge is that the processes of which our intuitions are the output are shaped by our evolutionary history. Thus, these processes have evolved to enhance our prospects of surviving and reproducing and are not connected with what we want our theories to represent (Singer 2005; see also de Lazari-Radek and Singer 2012).

Second, it could be the case that intuitions having property P is taken to suggest that they are unreliable, which means that they measure something inconsistently. According to this way of explicating the premise, intuitions can be valid but unreliable. Again, assuming that we are moral realists and that we want our intuitions to track moral facts, it could be that our intuitions track moral facts (i.e., are valid), but that they are easily distorted and thus that their measurement of moral facts often gets awry (i.e., track moral facts but do so unreliably). The second way of explicating the premise needs an account of a reliability threshold and evidence showing that intuitions fall short of this threshold (see e.g. Demaree-Cotton 2016; Mukerji 2018, 50; Sauer 2018, 75-76 for relevant discussions).

To illustrate, demographic variation and variation across situations in intuitions should suggest that intuitions are unreliable in the sense that, while they sometimes track what we want them to track, demographic and situational factors distort them. What distinguishes arguments targeting the reliability of intuitions from those targeting the validity of intuitions, is that the former relies on the premise that intuitions are influenced by irrelevant factors whereas the latter relies on the premise that they track irrelevant factors.\(^5\)

### 3.5 Is Experimental Philosophy a Game-Changer?

One may wonder what the novelty of experimental arguments is. Peter Singer (1974) argued long before the emergence of experimental philosophy that our intuitions are the products of cultural indoctrination and, therefore, not suited as data for moral theorizing. Thus, arguments following something like the scheme described in Section 3.1 have been put forward

---

\(^5\) It could be questioned whether (some of the explications of) the arguments from experimental philosophy are equally damaging to all metaethical positions. Sharon Street (2006, 2016; see also Kahane 2011, 112), for example, seems to think that evolutionary debunking argument are not damaging constructivist positions. Katia Vavova (2021, 720-721; see also Joyce 2016, 145-146), however, claims “Whatever our ethical or metaethical commitments, I suspect most of us believe that morality is about more than what is in our evolutionary interest”, thus, “one needn’t be a moral realist to be vulnerable to this attack.” In this thesis, I will grant that the question of whether intuitions are suited for serving as data for moral theorizing may depend on, for example, metaethical considerations.
independently of experimental philosophy. Moreover, both traditional and experimental philosophers are concerned with explaining our intuitions. They prompt intuitions about cases and try to determine what factors they respond to—i.e., they try to develop principles and theories that can account for our intuitions. What, then, is new? That is, what is the contribution of experimental philosophy?

First, while both traditional and experimental philosophers try to determine what factors our intuitions respond to, most traditional philosophers appear to have been working on the assumption that they respond to morally relevant factors and therefore sought to develop moral principles and theories to account for our intuitions. On the other hand, experimental philosophers do not, for most part, seem to proceed on this assumption. To a larger extent, they have developed psychological theories (that do not necessarily assume that our intuitions respond to morally relevant factors) to account for our intuitions (see Kahane 2013, 424-428).

Second, while both traditional and experimental philosophers are studying intuitions, they use different methods for collecting them. Traditional philosophers base their inquiries on their own intuitions. Thus, their inquiries can be conducted from their armchairs. Experimental philosophers study the intuitions of others—i.e., the participants in their experiments or those answering their surveys. Studying the intuitions of others allows experimental philosophers to study variation across demographic groups and situations, which is not something one can do from one’s armchair.

Moreover, experimental philosophers have tools that, for example, enable them to study which parts of their participants’ brains that are active when the participants make judgments. This allows them, to put it roughly, to examine whether participants’ judgments are driven by emotions or are the output of analytic thinking. One could suggest that traditional philosophers have been able to investigate whether their judgments are driven by emotions through the means of introspection. However, while the introspective study conducted by a traditional philosopher only produces evidence about the underpinnings of the philosopher’s judgments in the form of testimony, experimental studies provide evidence in the form of brain scans as well (which in turn can be used to validate the credibility of testimonial reports of the underpinnings of one’s judgments). The methods of experimental philosophers, then, provide access to a lot of information that previously has not been available to traditional philosophers (see also Königs 2020, for a similar argument).6

6 It is worth noting that some (experimental) philosophers use surveys to collect intuitions and then use these intuitions to test and develop moral principles and theories (see e.g. Lindauer and Barry 2017; Inoue et al. 2019; see also Carter, Peterson, and van Bezooijen 2016). What distinguishes this group of philosophers from traditional
Having noted these differences, let us now return to Singer’s arguments. To adequately justify their descriptive premises, that intuitions are the products of cultural indoctrination, he needs evidence of the kind provided by experimental philosophy studies. Thus, while traditional philosophers have put forward arguments similar to arguments from experimental philosophy, the arguments from experimental philosophy pose a more serious challenge as the descriptive premises are better supported by evidence (Alexander 2012, 28, 113; Kahane 2016, 299; Kumar and May 2019, 43-44).

3.6 Challenging the Method in Moral Philosophy

Having outlined in quite some detail the logic of experimental arguments, we may now take a closer look at the step from experimental arguments to experimental challenges—i.e., from arguing that intuitions lack initial credibility to arguing that a specific method (in this case, MRE) is flawed. Singer (2005, 348) accounts for this step in terms of a particular level and a general level:

[R]ecent 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. We should, with our current powers of reasoning and our rapidly changing circumstances, be able to do better than that.

philosophers is mainly the way of collecting intuitions. This begs the question; what is the better method of collecting intuitions? Of course, if it turns out that the method experimental philosophy researchers apply is better (for some reason), it could pose a challenge to traditional armchair methods, as it would suggest that philosophers need to get out of their armchair for collecting intuitions (Machery et al. 2013, 621). However, I will leave that kind of challenge aside in this thesis.

That said, Nobert Paulo (2019, 68), while criticizing the reliance on intuitions as data and MRE more generally, thinks that the evidence suggests that the influence of morally irrelevant factors on moral intuitions is “not as significant and widespread as the early critics seems to have suspected.” Thus, while strengthening the empirical premise of arguments against treating intuitions as data, experimental philosophy studies may also nuance assertions made by critics of reliance on concrete intuitions as data for moral theorizing.
Based on the findings of experimental philosophy studies we can formulate experimental arguments. We can use experimental arguments to construct experimental challenges. To move from an experimental argument to an experimental challenge, one must specify why the fact that our intuitions do not reflect morally relevant factors defeats a particular method. Singer accounts for the step as follows: Singer claims that it follows from the fact that our moral intuitions respond to morally irrelevant factors that the correct moral theory possibly conflict with all our intuitions (Singer 2005, 345-346). He argues that MRE appears to rule out this possibility. It does so “because it assumes that our moral intuitions are some kind of data from which we can learn what we ought to do” (Singer 2005, 346). In brief, according to Singer, experimental arguments pose a challenge to MRE because, according to MRE, our intuitions are some kind of data.

As I argued in Section 2, a common assertion is that moral intuitions track whatever our moral principles and theories are supposed to represent. Thus, we may account for the step from the experimental arguments to experimental challenges as follows: If relying on intuitions for testing the external validity of principles and theories is an essential feature of a method, and intuitions are unreliable (to the extent that they cannot serve this purpose) or invalid, then the method is flawed.

3.7 A Note on the Controversy of Experimental Arguments

The findings of (some) experimental philosophy studies that purport to establish that our intuitions have some problematic properties are contested. Some of the findings have not been replicated (see e.g. Seyedsayamdost 2015a; Seyedsayamdost 2015b; see also Johnson, Cheung, and Donnellan 2014; and Schnall 2014, for commentary and a rejoinder on Johnson, Cheung, and Donnellan 2014),8 and the methodology of (some of) the studies have been criticized (Bergenholtz, Busch, and Praèm 2021; Cullen 2010; Thompson 2022; Woolfolk 2013). Some have objected that experimental philosophy studies lack relevance because they mainly concern the intuitions of non-philosophers. The assumption is that philosophers’ intuitions are better (e.g., in the sense of being more reliable) than the intuitions of non-philosophers (see e.g.

8 But see Cova et al. (2021). They argue that in general, the replication rate for experimental philosophy studies is quite high. They successfully replicated about 70% of the studies in their sample (N = 40). However, the replication rates varied with the type of effect studied by experimental philosophers. The average replication rate for demographic effects, which is the type of effect that Seyedsayamdost tries to replicate, is only 25% (see also Knobe (2019, 2021) for an argument that the current available evidence on demographic effects suggests that intuitions are quite stable across demographic groups). Note also that the studies that Seyedsayamdost tries to replicate are not included in Cova et al.’s sample. See also a special issue in Review of Philosophy and Psychology for more on replicability in experimental philosophy and cognitive science more generally (Strickland and De Cruz 2021).
Andow 2015a; Grundmann 2010; Horvath 2010; Ludwig 2007). Some have even questioned whether experimental philosophers capture what philosophers call intuitions in their studies at all (see e.g. Bengson 2013). Moreover, it has been suggested that some of the properties are not as problematic as some critics claim and do not warrant depriving intuitions of their (initial) credibility (Enoch 2010). More generally, there is an ongoing discussion regarding what inferences one can draw from the findings of experimental philosophy studies (Berker 2009; Königs 2018; Lott 2016; see also Locke 2020).

In light of these considerations, the current thesis seems to grant too much with regard to the argument from experimental philosophy. Thus, it seems to start off from a rather controversial starting point. How interesting an investigation is that proceeds from such a controversial starting point might be questioned. However, several considerations suggest this endeavor is worthwhile.

First, although some findings have not been replicated, others have. Experimental philosophy is a recent movement, and more studies are needed to establish what properties our intuitions have. It could be argued that it is premature to assume that experimental philosophy is unable to establish what properties our intuitions have.

Second, although these arguments are controversial, philosophers are advancing such arguments; thus, responses are warranted. One way of responding to them is to criticize the arguments from experimental philosophy. Another way of responding to these arguments is to examine what implications they might have. Nicholas Sturgeon (1986, 115) suggests that “any critical response to a philosophical position can be classified either as an ‘Oh yeah?’ or a ‘So what?’” An “Oh yeah” response amounts to questioning the truth of the philosophical position, while a “So what?” response amounts to questioning the significance or implications of the position. Thus, in Sturgeon’s terms, this thesis primarily explores a “So what?” response to the arguments from experimental philosophy.

Although some experimental philosophers have, in response to this objection, conducted studies with philosophers as participants (Schwitzgebel and Cushman 2012, 2015; Tobia, Buckwalter, and Stich 2013; Tobia, Chapman, and Stich 2013; Löhr 2019; Horvath, Meyer, and Wiegmann 2020; Horvath and Wiegmann 2021) arguing that philosophers’ intuitions are no more reliable than nonphilosophers’, this issue is far from settled. In response to such studies, it has been argued that the expertise of philosophers does not consist in having more reliable intuitions, but rather in being better at identifying which intuitions are reliable (see e.g Kauppinen 2018, 65-67; see also Nado 2015, 1036; Drożdżowicz 2018; Egler and Ross 2018) or that the studies do not examine philosophers’ intuitions in “the theoretical and practical context in which such expertise is ordinarily exercised” (Kauppinen 2018, 50, italics removed; see also De Cruz 2015).

Note that some of these arguments target the use of methods from neuroscience in particular.
Finally, even if it should turn out that experimental arguments are fundamentally flawed (e.g., because intuitions do not have the problematic properties), this thesis contributes to a debate about a more general question: what role do intuitions play in moral inquiry and to what extent does MRE rely on intuitions as data?
So far, I have reviewed the literature on moral intuitions and MRE and outlined the challenge from experimental philosophy. The impression one gets from reading the literature on moral intuitions and MRE is that moral intuitions play an important role in moral inquiry and figures centrally in MRE. Consequently, the challenge from experimental philosophy seems to shake the methodological foundations of political philosophy and moral philosophy more generally. However, before I conclude the literature review, one more strand of literature relevant to this thesis needs to be reviewed. Lately, several philosophers have challenged the picture of philosophy outlined in Section 2. Some philosophers, such as Herman Cappelen (2012) and Max Deutsch (2009, 2010, 2015a, b), argue that philosophers do not treat concrete intuitions as conferring any positive epistemic status (see also Horvath 2022). Following Cappelen (2012), I will refer to the claim that intuitions are treated as conferring some positive epistemic status as “Centrality.” If Centrality is false, as mentioned in the introduction, it would seem to relieve philosophers of the challenge from experimental philosophy. Following Jennifer Nado (2016), I will refer to those who reject Centrality as “intuition-deniers.”

In this section, I give an overview of the Centrality debate. I outline some of the arguments pressed against Centrality: (i) that philosophers’ use of intuition-terminology does not support Centrality, (ii) that salient facts about the philosophical practice suggest that Centrality is false,

---

1 In the current debate, it has been common to explicate the positive epistemic status in terms of “evidence”. In this section, Section 6, 7, and Part 2 of the thesis, I will follow this convention and ask and inquiry whether intuitions are treated as evidence. However, I do not think that matters much substantially. I think that most of the philosophers arguing that intuitions are not treated as evidence would deny that intuitions confer any positive epistemic status (although they might be useful in other contexts, such as the context of discovery (Cappelen 2012, 230)). At least, I think if their arguments are successful, the arguments show not only that intuitions are not treated as evidence, but that intuitions are not conferring any positive epistemic status. Similarly, I do not think that my arguments commit me to a specific explication of the positive epistemic status but is compatible with intuitions conferring any positive epistemic status.
and (iii) that when we examine how philosophers philosophize, it is evident that Centrality is false. Moreover, I also outline some of the responses to these arguments by friends of Centrality.

### 4.1 Intuition-Terminology

One way in which Centrality has been challenged is by attacking the assumption that philosophers using “intuitive” and cognate terms (henceforth, “intuition-terminology”) to refer to key premises in their arguments suggests Centrality is true (Cappelen 2012; Dorr 2010; see also Andow 2017a). I will refer to this challenge as the “terminology challenge.” Let me stress that the intuition-deniers pursuing this strategy do not deny that philosophers use such terms. Rather, they claim that the way philosophers use such terms does not support Centrality. Cappelen, for example, looks at both how intuition-terminology is used in ordinary English and philosophers’ special technical use of such terminology. He argues that neither use of intuition-terminology supports Centrality. In ordinary English, intuition-terminology functions as hedges rather than indicating that the speaker relies on a special kind of evidence or evidential source (Cappelen 2012, ch. 2). Philosophers’ use of intuition-terminology is defective (Cappelen 2012, ch. 3). He claims that there is no agreed-upon definition of intuition or set of paradigms of intuition, and philosophers disagree about the theoretical role of intuitions. This, he argues, makes it unclear how to interpret philosophers’ use of intuition-terminology.

However, he suggests some “charitable” reinterpretations of philosophers’ use of intuition-terminology (Cappelen 2012, ch. 4). For example, in many cases one can simply remove the intuition-terminology. Such removal, he claims, will not lead to argumentative loss. This suggests that the intuition-terminology plays no argumentative role. He argues that some intuition-terminology can be reinterpreted as hedge-terms. That is, terms which function is to weaken the speaker’s commitment to a claim. Another suggestion he makes is that terms such as “intuitively” may mean something close to “pre-theoretically.” This could, for instance, suggest that such terminology is used to indicate that a claim is in the common ground; a claim that the participants in a debate already agree on and are committed to. None of Cappelen’s reinterpretation strategies, however, yield support to Centrality.

Underlying the terminology challenge, then, are the following observations: the ordinary English use of intuition-terminology does not support Centrality, philosophers’ use of intuition-terminology...
terminology is defective, and charitable reinterpretations of philosophers’ use of intuition-terminology do not support Centrality.

There are several responses to the terminology challenge. Jonathan Weinberg (2016, 289) has pointed out that we can “locate a fairly broad agreement across the profession as to where intuitions are supposed to be found, even while facing this broad dissensus as to what intuitions are supposed to be.” Moreover, he notes that

even philosophers who operate with radically different professed accounts about what makes the appeal to intuitions legitimate, are nonetheless perfectly happy to share cases back and forth with each other, without batting an eye—and without drawing any methodological consequences from their meta-level disagreement to their first-order debates. (Weinberg 2016, 295)

Weinberg, then, argues that although philosophers provide different substantial accounts of intuitions, these accounts are all accounts trying to capture the same phenomena. That is, philosophers do not disagree about what “intuition” refers to. Rather, they disagree about how to account for the phenomena referred to as “intuition.”

Similarly, Michael Devitt (2015, 677-680) claims that when considering whether Centrality is true or not, the nature of intuition is beside the point. Devitt agrees that differences in substantial accounts of intuitions are largely irrelevant for the task of interpreting philosophers’ use of intuition-terminology. To illustrate, from the fact that some philosophers think of intuitions as sui generis mental states while others think of them as beliefs or inclinations to believe, it does not necessarily follow that they are referring to different phenomena when using terms like “intuition.” For example, they may all agree that when presented with the pond case, a mental state occurs in us with the content that we should help the small child and that “intuition” refers to that mental state. They may simply disagree on the correct account of that mental state.

Weinberg and Devitt, then, claim that Cappelen’s conclusion about the defectiveness of philosophers’ use of intuition-terminology is, at least in part, reached by wrongly inferring that philosophers disagree about what phenomena intuition-terminology refers to from the fact that philosophers disagree about the nature of intuition. In Devitt’s terms, Cappelen fails to satisfactorily distinguish between what philosophers take to be the nature of intuition and what philosophers take to be the meaning of terms like “intuition.” What philosophers mean by the term “intuition” is a mental state (or the content of such a mental state) formed non-inferentially,
which is, he claims, what “intuition” means in ordinary English (Devitt 2015, 673-684, see also Section 2.1.1 of this thesis).

What Devitt refers to as “the meaning of ‘intuition’” (i.e., a mental state formed non-inferentially, or the content of such a mental state), I take it, can be understood as a thin account of intuition that aims at describing the relevant phenomenon without making substantial claims about the nature of it. Such accounts are compatible with thicker accounts of intuition making substantial claims about the nature of intuitions. So while philosophers provide different accounts of the nature of intuition, there is still broad agreement about the thin account of intuition (see Section 2.1.1).

Bengson’s (2014) response to Cappelen is slightly different from Devitt’s. Bengson argues that Cappelen fails to consider a third use of intuition-terminology, which he calls “discriminative use” of intuition-terminology. Bengson (2014, 559) suggests that there is a specific use of ‘intuition’ and cognate terms that, while rooted in natural language, is more discriminating (i.e., it has a narrower extension) than what is found in casual ordinary discourse, but which is not merely “special technical”—though it might be put to a theoretical purpose.[…] In effect, even if the relevant use is somehow inappropriate or odd (e.g., “deviant”, not “core”, not “central”; cf. 27, 42) from the perspective of ordinary discourse, it is in a sense already out there, not invented by contemporary philosophers but latent in natural language, although it may require extensive investigation, by theorists of intuition, to reflectively disentangle and discern.

(Nota removed)

Like Devitt, Bengson thinks that philosophers’ use of intuition-terminology is not embedded with theoretical commitments. However, Bengson does not think this way of using intuition-terminology completely overlaps with how such terminology is used in ordinary English as it is more discriminative. By that, he does “not intend to suggest that ordinary use fails to discriminate, but rather that the relevant use discriminates where the ordinary use does not” (Bengson 2014, 559n8). That said, I think Bengson’s response is in the same spirit as Devitt’s response, as the discriminative use of intuition-terminology is not completely detached from the use of intuition-terminology in ordinary English. As Bengson remarks, the discriminative use of intuition-terminology “is in a sense already out there” in “ordinary discourse.”

To summarize the discussion so far, intuition-deniers have argued that the use of intuition-terminology does not indicate that intuitions are treated as evidence and that intuition-terminology might not even be used to designate intuitions at all. In response, it has been objected that this argument misconstrues how philosophers use intuition-terminology.

42
Philosophers’ use of such terminology is similar to how it is used in ordinary English, that is, to designate a mental state (or the content of the mental state) of how one is intellectually appeared-to.

While I find these responses plausible, I think they fall somewhat short of responding to the arguments from the intuition-deniers. The claim made by the intuition-deniers is that the presence of intuition-terminology does not indicate that intuitions are treated as evidence. While the responses succeed in arguing that the presence of intuition-terminology is good evidence of the presence of the relevant mental state (or that some claim is the content of the relevant mental state), the arguments fall short of establishing that the presence of intuition-terminology shows that that very mental state is treated as evidence. In other words, I think the arguments concerning how philosophers use intuition-terminology do not provide positive evidence of the reliance intuitions as evidence, but merely suggest that intuition-terminology is used to designate the relevant mental state.

### 4.2 Explaining Salient Facts about the Philosophical Practice

Some intuition-deniers have argued that some (putative) facts about the philosophical practice pressure Centrality. Bernard Molyneux (2014, 446), for example, has claimed that the following phenomena count against Centrality:

(A) **Insufficiency**: The propositions that intuition putatively supports are treated as having a degree and kind of certainty and justification that they could not have got from being intuited [...].

(B) **Liberal applicability**: Intuitions influence us in ways we cannot explain by supposing we treat them as evidence [...].

(C) **Anomalous inadmissibility**: Certain strong intuitions that persuade us of their contents are treated as inadmissible in the context of justification [...]. (Italics in original)

On the other hand, friends of Centrality have argued that some (putative) facts about the philosophical practice are better explained by Centrality than any explanation denying Centrality. According to Nevin Climenhaga (2018, 97), the following are examples of such facts:

- **Fact 1**: Philosophers frequently believe propositions they find intuitive.
- **Fact 2**: Philosophers offer error theories for intuitions that they reject.

---

3 See also Andow (2017b) for a reply to Molyneux.
Fact 3: Philosophers are more confident in rejecting (accepting) theories when they have intuitions about diverse cases opposing (supporting) them.\(^4\)

Thus, there is disagreement regarding whether central features of philosophers’ practice favor the acceptance or rejection of Centrality. It is worth noting that the articles mentioned in this section, although drawing on examples from the philosophical articles on first-order issues, does not engage in systematic case studies of philosophical articles to support their arguments. In the next section, I review some case studies investigating Centrality and the discussion sparked by those case studies.

### 4.3 Case Studies

Intuition-deniers have done case studies of influential philosophical articles. On the basis of these case studies, they have argued that there is no reason to believe that intuitions are treated as conferring any positive epistemic status. The most extensive case studies so far have been the case studies conducted by Herman Cappelen (2012, Part 2) and Max Deutsch (see in particular 2015b). This strategy for arguing for the falsity of Centrality consists of the following steps: first, identify some features that we would expect to be present in a text if intuitions were treated as conferring any positive epistemic status, and second, look for these features in philosophical articles. I will in the following outline some of the features identified by intuition-deniers and the main findings of their case studies. In particular, these case studies have been concerned with philosophers’ discussion of cases. More precisely, these case studies have been investigating whether the intuitiveness of claims about cases is treated as conferring some positive epistemic status. Thus, these case studies have investigated whether philosophers treat concrete intuitions as conferring positive epistemic status.\(^5\)

One suggestion is that a characteristic of “appeals to the intuitive” (Cappelen 2012, 111) is that “an intuitive judgment has a characteristic phenomenology” (Cappelen 2012, 112). Cappelen applies this strategy in his case study. Cappelen illustrates what he has in mind when he writes “characteristic phenomenology” by citing Plating (1993, 105-106, cited in Cappelen 2012, 9, 101 and 2014, 588): “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 proposition is true” (my italics). Therefore, one strategy is to search for

\(^4\) See also Horvath (2022, 26-27) for a discussion of Climenhaga’s argument.

\(^5\) It is worth noting that Cappelen (2012, 196), for example, is open to the possibility that mid-level intuitions and abstract theoretical intuitions could be treated as conferring positive epistemic status. Deutsch (2015b, ch. 4, 5), on the other hand, is more skeptical to that possibility.
claims accompanied by this kind of phenomenology. There are two challenges pertaining to whether a claim has this characteristic. First, one needs a way of determining that a claim is accompanied by this characteristic phenomenology. Second, one needs a way of determining whether a claim is treated as conferring positive epistemic status because it is accompanied by this kind of phenomenology. Cappelen (2012, 118) suggests that one may overcome these challenges by looking for authors explicitly stating that a claim is accompanied with (or missing) the relevant kind of phenomenology that is treated as conferring (or not conferring) positive epistemic status to, for example, a principle. However, Cappelen does not find any such statements in the articles he investigates.

Another assumed characteristic of appeals to a claim’s intuitiveness (as the source of that claim’s positive epistemic status) is that the claim is not supported by arguments (Cappelen 2012, 121; Deutsch 2017, 429). That is, since the claim is granted positive epistemic status by virtue of its intuitiveness, there is no need to support it with arguments. However, according to the intuition-deniers, when a claim is supported by arguments, we have no reason to believe that the intuitiveness of the claim is conferring any positive epistemic status onto the claim. That is, there is no need to assume that the intuitiveness of the claim is doing any epistemic work. Both Cappelen and Deutsch argue that philosophers’ claims about cases are typically supported by arguments.

Moreover, the intuition-deniers argue that if a claim is “independent of the answers to a certain set of contextually determined questions, Q,” where Q typically “are the questions under discussion in the speech context” (Cappelen 2012, 68), the lack of arguments in support of the claim would not amount to evidence that the claim is treated as having positive epistemic status due to its intuitiveness. Rather, a more plausible explanation is that the claim “is in the common ground among the participants in the debate,” that is, a claim that is endorsed by the participants in the debate and that they already are committed to (Cappelen 2012, 153, see also pp. 154-156).

While Cappelen (2012, 153) finds a few claims that he thinks are asserted to be in the common ground, the general picture that emerges from Cappelen’s and Deutsch’s case studies is that philosophers argue for their claims. That is, they do not find claims that are not either argued for or asserted to be in the common ground.

To illustrate, consider Deutsch’s (2015b, 80) interpretation of Gettier’s (1963) argument against the justified true belief (JTB) theory of knowledge. According to Deutsch (2015b, 41), Gettier’s argument targets the following generalization implied by the JTB theory: “For every subject, S, and every proposition, p, if S justifiably and truly believes that p, then S knows that
The JTB theory is rejected on the basis of a judgment about the ten coins case. In this case, Gettier (1963, 122) asks us to imagine that Smith and Jones have applied for some job. We are also asked to suppose that Smith has strong evidence for proposition (d): “Jones is the man who will get the job, and Jones has ten coins in his pocket.” Based on (d), Smith believes that (e): “The man who will get the job has ten coins in his pocket.” Thus, “[i]n this case, Smith is clearly justified in believing that (e) is true.” However, “unknown to Smith, he himself, not Jones, will get the job,” and “unknown to Smith, he himself has ten coins in his pocket.” While the generalization of JTB holds, Gettier’s judgment is that Smith does not know that (e). This judgment, Deutsch (2015b, 81) claims, is supported by an argument, and cites the following paragraph as evidence:

it is [...] 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 coins in Jone’s pocket, whom he falsely believes to be the man who will get the job. (Gettier 1963, 122, italics in original)

Roughly, then, Deutsch seems to hold that Gettier’s refutation of JTB has the following form: JTB is false because it implies that Smith knows that I. However, Smith does not know that I. The claim that Smith does not know that I is supported by an argument based on something like the assumption that “knowledge cannot be inferred from false beliefs.” Cappelen and Deutsch, then, are not suggesting or presupposing that philosophers are mainly interested in the cases themselves. Rather, the idea is that philosophers argue for case verdicts and, in turn, make inferences about general philosophical theories based on those case verdicts. So the function of philosophical cases, according to this view, is, for example, “to raise a range of questions about some philosophically significant feature of the world” (Cappelen 2012, 188). The answers to these questions, in turn, inform a more general philosophical question.

A third feature of appeals to the intuitiveness of a claim, which is considered by Cappelen (2012, 113), is that the claim “is justified solely by the subjects’ conceptual or linguistic competence…” According the account of intuition underlying this suggestion, intuitions are the product of our conceptual competence. Here are some examples of strategies Cappelen (2012, 126-127) suggests would help us determine whether an author is appealing to the intuitiveness of a claim (on this account of intuition): (i) consider whether the author explicitly claims that a judgment is based solely on conceptual competence, (ii) consider whether the judgment in question is of a necessary truth, and (iii) consider whether a candidate claim is like those of
alleged paradigms of conceptually justified claims rather than claims that are not so justified. In his study, Cappelen does not find any statements like (i) or claims like (ii) and (iii), which he takes to be evidence against the intuitiveness of the claims conferring any positive epistemic status.

In sum, the intuition-deniers believe that the absence of these features in philosophical articles is evidence against the intuitiveness of claims conferring positive epistemic status. However, these strategies have been challenged by friends of Centrality. To the first feature, i.e., that intuitions have a characteristic phenomenology, it has been objected that, for example, intuitions do not have a “characteristic” or “special” phenomenology (Boghossian 2014, 370; see also Brogaard 2014, 388; Richard 2014, 404; Weatherson 2014, 519) and that, if intuitions are accompanied by any phenomenological state, it is a more mundane state that even the intuition-deniers seem to admit to be aware of (Climenhaga 2018, 70-71, 79).

To the third feature, it has been objected that the account of intuition according to which intuitions stem from conceptual competence alone is controversial, in particular with regard to normative and moral intuitions (Chalmers 2014, 536; see also Boghossian 2014, 372; DePaul 2006, 613; Richard 2014; Sosa 2007, 100; Weatherson 2014, 519). Moreover, even if it were the case “that the essential defining feature of intuition is that it is based in conceptual competence; it in no way follows that a philosopher who relies on intuition will recognize this—not even if she explicitly believes herself to be relying on intuition…” (Nado 2016, 794). Thus, the failure of the strategies pertaining to this feature is no surprise.

Regarding the observation that philosophers argue for their claims, it has been argued that philosophers arguing for a claim is compatible with philosophers treating the intuitiveness of that very claim as conferring positive epistemic status (see e.g. Bengson 2014, 572-574; Brown 2017, 197-199; Chalmers 2014, 539-541; Weatherson 2014, 526, 528; Weinberg 2014, 546; see also Devitt 2015, 694-695). Moreover, it has been argued that some of the arguments that Cappelen and Deutsch assume are arguments for claims, in fact, are abductive arguments in which the claim figure as the starting point. When a claim, c, figures as a starting point for an abductive argument, c supports the conclusion of that argument. Thus, in that case, the direction of support is the opposite of what Cappelen and Deutsch assume, which in turn means that the presence of the argument—i.e., the abductive argument—is not evidence against the intuitiveness of the claim is not conferring positive epistemic status (Colaço and Machery 2017, 413-415; Machery 2017, 179-180; Nado 2016, 795; 2017, 395-396; see also Bengson 2014, 572; Tobia 2015).
To these objections, the intuitions-deniers have responded, first, that arguing that a particular claim being argued for is compatible with the intuitiveness of that very claim conferring positive epistemic status onto that claim will not do. Merely showing that it is possible that the intuitiveness of the claim is conferring positive epistemic status is not positive evidence that the intuitiveness of the claim is conferring some positive epistemic status. Moreover, that an argument is provided in support for a claim suggests that there is no need for the intuitiveness of the claim to confer any positive epistemic status onto the claim (Cappelen 2014a, 599-600).

Moreover, the intuition-deniers have argued that in the cases where an argument is open to being read as an abductive argument where the claim about a case is treated as the starting point, the correct reading is that the direction of support goes in both directions—that is, from the claim about the case to the theoretical account and from the theoretical account to the claim about the case (Cappelen 2012, 123). Hence, the argument is not “fully” abductive (Deutsch 2017, 438). The assumption among the intuition-deniers seems to be that it is only when the argument is fully abductive, and the claim about a case is treated as a starting point, that we need something like the intuitiveness of the claim to make sense of the argument. As long as we do not need to assume that the intuitiveness of the claim is doing any epistemic work for the argument to work, we should refrain from assuming so (see also Horvath 2022; Horvath and Koch 2021, 6).

4.4 Taking Stock

So where does this leave us, then? As the discussion above implies, according to the intuition-deniers, the burden of evidence is on friends of Centrality to show that intuitions are treated as conferring some positive epistemic status. However, the claim that the burden of evidence is on friends of Centrality has been questioned (see e.g. Nado 2016).

Moreover, the evidence provided so far seems to be inconclusive. It is contested whether the intuition-deniers have provided evidence incompatible with Centrality. Moreover, it is contested whether (and outright rejected by the intuition-deniers that) friends of Centrality have provided positive evidence for Centrality is contested. More disturbingly, the indeterminacy of the evidence from the case studies, and the discussion of these studies, suggest that we lack a proper strategy for testing Centrality.

Thus, according to this diagnosis, there are two ways of advancing the debate further. Either (i) successfully push the burden of evidence onto one of the sides of the debate or (ii) take on the burden of evidence. I attempt the latter.
Picking Up the Threads and Outlining the Contribution of the Thesis

The relevant literature having been reviewed, it is time to pick up the threads and explain how this thesis contributes to the literature reviewed. To do so, it would help revisit the research questions investigated by this thesis:

Overarching question: Does the challenge from experimental philosophy succeed?
I. Do contemporary analytic political philosophers treat the criterion of intuitive fit as a theoretical desideratum?
II. Does MRE presuppose the criterion of intuitive fit?

I defined the challenge from experimental philosophy as the claim that arguments from experimental philosophy—i.e., roughly, the argument that the findings from experimental philosophy show that intuitions should not be treated as conferring any positive epistemic status and hence that the criterion of intuitive fit is flawed—show that philosophers must radically change their methods. Moreover, I distinguished between two interpretations of the challenge: (1) a challenge targeting the ideal of MRE, and (2) a challenge targeting how most proponents of MRE currently apply the method.

To clarify the purpose of the current thesis and to explain how the two ways of understanding the challenge relate to each other, I suggested that it is helpful to consider the following claims:

A. Most political philosophers apply MRE
B. MRE can only be a valid method if intuitions can be treated as data
C. According to the findings of experimental philosophy studies, intuitions are not suited for serving as data for moral theorizing

D. Most political philosophers treat intuitions as conferring positive epistemic status

To reiterate, the first formulation presupposes C and B. The second formulation presupposes A, D and C. This thesis grants A, and to some extent C, and primarily investigates B and D, hence questions (I)—Do contemporary analytic philosophers treat the criterion of intuitive fit as a theoretical desideratum? —and (II)—Does MRE presuppose the criterion of intuitive fit? Answering (I) is crucial for assessing whether the version of the challenge targeting the current implementation of MRE succeeds and answering (II) is crucial for determining whether the challenge targeting the ideal of MRE succeeds.

In the review of the literature concerning D (Section 4), I argued that there is currently a lack of a proper method for testing Centrality. In investigating research question (I), this thesis aims to provide such a method and thereby contribute to the debate on whether philosophers treat intuitions as conferring some positive epistemic status. Moreover, since this thesis is concerned with the method in political philosophy in particular, it is worth noting that when it comes to case studies of the kind described in Section 4.3, very few of the articles studied come from political philosophy and neighboring disciplines. Thus, another important contribution of this thesis is that it remedies the lack of case studies of articles from political philosophy. Since there could be variation in what methods are applied across philosophical disciplines, it is important to investigate cases from political philosophy.

The thesis also makes important contribution to the current literature in investigating (II). First, it introduces several distinctions that allows for a more fine-tuned discussion of the challenge from experimental philosophy. The thesis not only distinguishes between narrow and wide MRE, but also distinguishes between different interpretations of MRE within the category of wide MRE. Moreover, as mentioned above, it distinguishes between different readings of the challenge from experimental philosophy. Second, building on the discussion of the challenge from experimental philosophy, the thesis offers an explication of what the distinctive features of MRE is.

Looking ahead, I argue in this thesis that D is true and argue that B (if we consider the ideal of MRE) is false. As pointed out in the introduction, this seems to be an inconsistent position. Accepting that most philosophers apply MRE (A) and rejecting that MRE presupposes that intuitions are treated as conferring some positive epistemic status (B) on the one hand, seems to be inconsistent with accepting that most philosophers treat intuitions as conferring some
positive epistemic status (D) on the other hand. It might seem, then, that the thesis defends MRE from the challenge from experimental philosophy only by ending up accepting an inconsistent position—which does not amount to a defense at all!

However, this impression is wrong. A, not-B, and D are consistent because while MRE is compatible with treating intuitions as data, it is not an essential feature of MRE that intuitions are so treated. The way most proponents of MRE apply the method appears to involve treating intuitions as data. What follows, then, is that the challenge from experimental philosophy does not undermine particular interpretations of the ideal of MRE because they do not presuppose the criterion of intuitive fit. However, the arguments from experimental philosophy apply to MRE as it is currently implemented by many of its adherents. More specifically, how MRE is currently applied seems to involve treating some background assumptions as fixed. These background assumptions include, for example, the assumption that concrete intuitions can be treated as data. Although it is treated as fixed, this assumption is not essential for MRE. The assumptions in question are not treated as fixed because it is believed that the assumption has a special epistemic status or because it is assumed that that issue has a priority in the sense that it must be decided prior to engaging in substantive moral inquiries. It might simply be treated as fixed out of convenience. As soon as the reasons speaking against this assumption become decisive, MRE allows for rejecting the assumption. Revising this assumption, then, would not demand the rejection of MRE as such.

In conclusion, then, I answer (I) in the affirmative. I answer (II) in the negative if we consider (particular interpretations of) the ideal of MRE and in the affirmative if we consider the current implementation of MRE. That is, when the challenge from experimental philosophy is interpreted as a challenge targeting the ideal of MRE, the challenge fails. However, I argue that when the challenge is interpreted as targeting most philosophers’ current implementation of MRE, and we grant A and C, it succeeds.
In this section, I discuss ethical considerations concerning the thesis and its relationship to the philosophy of science.

6.1 Ethical Considerations

There are several ethical considerations that apply to researchers.¹ This thesis does not process personal data; thus, ethical considerations pertaining to the processing of personal data do not apply to this thesis. However, there are other considerations that apply. For example, The National Committee for Research Ethics in the Social Sciences and the Humanities (NESH) mentions in Section D (The research community) of their guidelines considerations such as good citation practice, scientific integrity, and impartiality. This thesis aspires to fully satisfy these standards. Moreover, Section A of the NESH guidelines accounts for what is required for research to be considered responsible. For example, it states that “[g]reat demands are placed on the justifications of the researchers for their choice of questions, methods and analytical perspectives, and also on the quality of the documentation used to support conclusions, so that preconceived notions and unwitting opinions have minimal influence on the research.”² The next subsection concerns the methodological requirements mentioned in the quote.


² Although this quote is not from the latest edition of the guidelines (which, at the point of writing, yet have to be published in English), the importance of the mentioned methodological requirements are also stressed in the latest edition of the guidelines (published in Norwegian): [https://www.forskningsetikk.no/retningslinjer/hum-sam/forskningsetiske-retningslinjer-for-samfunnsvitenskap-og-humaniora/](https://www.forskningsetikk.no/retningslinjer/hum-sam/forskningsetiske-retningslinjer-for-samfunnsvitenskap-og-humaniora/) [Accessed 17.05.2022].
6.2 Relationship to the Philosophy of Science

The kinds of inquiries required to respectively investigate research questions (I) and (II) differ in important ways. The inquiry of (I) is mainly descriptive. The aim is to investigate what political philosophers do when they philosophize. However, the inquiry of (II) is philosophical. For these reasons, there are different questions from the philosophy of science that relate to the inquiry of (I) on the one hand and to the inquiry (II) on the other. Therefore, I will treat these inquiries separately in this section.

6.2.1 The Inquiry of Research Question (I)

The inquiry of (I) is descriptive not only in the sense that it is non-evaluative (as opposed to normative inquiry), but it is also in the sense that it aims at answering a *what* question about a phenomenon (as opposed to inquiries into causal relationships, which aim at answering *why* questions) (Gerring 2012, 722-723). *What* questions about a phenomenon need not be questions starting or including “what.” They can also ask, for example, “when” or “whom” (Gerring 2012, 722). In this sense, (I) pertains to *what* philosophers do, that is, whether they treat the criterion of intuitive fit as a theoretical desideratum.

It would be helpful to think of the inquiry of research question (I) as making inferences at two levels. First, inferences are made about a particular philosopher’s method based on a set of observations. Second, inferences are made about the method of a population of philosophers based on a sample of philosophers. Moreover, since (I) is descriptive (as opposed to causal), and it is suggested that a methodological issue concerning descriptive questions is the problem of labelling (Gerring 2012, 738), I will briefly discuss whether this problem applies in the case of (I). I will start by describing the method I will apply in my inquiry of (I).

6.2.1.1 Method

My method for investigating whether intuitions are treated as evidence consists primarily of applying an analytical framework that I develop in Article 1. According to this framework, to determine whether claims about cases are treated as evidence because of their intuitiveness, we must start by considering whether the case verdict is treated as evidence. We have reason to believe that a case verdict is treated as such if it is treated as a starting point for an abductive argument. The analytical framework offers diagnostics for identifying whether the case verdict is so treated. The next and final step of the framework is to determine whether case verdicts are

---

3 As in Articles 1 and 2, I now follow the convention of the Centrality debate of speaking of intuitions being treated as “evidence” (see Footnote 1, Section 4).
treated as starting points for abductive arguments because of their intuitiveness. If that is the case, we have reason to believe that the case verdict is treated as evidence because of its intuitiveness.

I use this framework to analyze several articles, namely, Dworkin’s “What is Equality? Part I” and Cohen’s “On the Currency of Egalitarian Justice” (Article 1, Section 8), as well as Thomson’s “The Trolley Problem” (Article 2, Section 9). Moreover, in Article 2, I also apply two additional strategies: I look at what the philosophers whose article I analyze claims to be her evidence and what she writes about her method.

6.2.1.2 Inferences Concerning an Author's Method

There are several available strategies for determining a particular philosopher’s method:

a. Consider what the philosopher says about her method.

b. Consider what the philosopher does when she philosophizes.

c. Consider what is perceived by the philosopher’s interlocutors to be her method.

One could also use a combination of these strategies. Looking ahead, my approach will be a combination of these strategies. To better understand why it makes sense to combine several strategies, it is helpful to first consider each strategy on its own and describe some of the problems that pertains to them when they are pursued individually.

First, (a) seems, at face value, to be a reasonable strategy. If you want to know what a person is doing, just ask the person directly. The person should be in a pretty good epistemic position with regard to knowing what she is doing. However, it has been objected that philosophers may have false second-order beliefs about their methods, in which case reports from the philosophers themselves would be invalid (Cappelen 2012, 164; Deutsch 2015b, 39; Molyneux 2014, 445). The worry is that since Centrality figures in many philosophers’ self-conception (for some reason that allegedly has nothing to do with how philosophers actually philosophize (Cappelen 2012, 21-23; Deutsch 2015b, xvii)), philosophers would believe that they treat intuitions as evidence even if Centrality were false. Thus, some claim that analyses under strategy (b) carry much more epistemic weight than those under strategy (a) (Horvath 2022, 30; Horvath and Koch 2021, 6).

However, strategy (b) also faces problems. Crucially, as Section 4.3 demonstrated, the findings of currently available studies that use strategy (b) are contested. The participants in the Centrality debate disagree on what inferences are warranted regarding the role of intuitions in
philosophy from the observations made in the studies. To illustrate, let us consider two observations: a philosopher makes a claim about a particular case, and this philosopher embraces a theoretical account that is consistent with the case verdict. What is under discussion is what we should infer about the relation of support between the claim and the theoretical account. The candidate inferences are as follows:

(i) The claim about the case supports the theoretical account  
(ii) The theoretical account supports the claim about the case  
(iii) The direction of support goes in both directions

That rather different conclusions can be inferred from the observation illustrates that one needs to be cautious about what inferences one makes when engaging in strategy (b). More specifically, in this case, we are dealing with the problem of underdetermination. The current observations do not favor any of the three claims.

Moreover, even if we are able to show that (i)—the claim about the case supports the theoretical account—is the case, it is still an open question whether the claim supports the theoretical account because of its intuitiveness. According to the intuition deniers, the claim may be treated as such because it is in the common ground for some extra-intuitive reason.

Strategy (c), I assume, is prone to the same objections as strategy (a), at least to some extent. Because Centrality figures in many philosophers’ self-conception, philosophers may believe that other philosophers philosophize according to Centrality even when this is not the case.

To investigate in more detail the problems concerning the strategies, let us consider a couple of scenarios. Let us assume that a philosopher, S, has written an article defending C and that P is S’ interlocutor. Now, consider the following scenarios:

**Scenario A:** S thinks she is treating both argument A and the intuitiveness of C, I, as evidence for C in the article she has written. In fact, this what S is actually doing. However, P does not think that S treats I as evidence. Rather, P thinks that S treats A, and only A, as evidence for C. Thus, after having read S’ paper, P accepts C because of A.

**Scenario B:** S thinks she is treating A, and only A, as evidence for C. In fact, this is what S is actually doing. However, P does not think that S treats A as evidence for C. Rather, P thinks that S treats I, and only I, as evidence for C. Thus, after having read S’ article, P accepts C because of I, and only I.
The question is, then, do the philosophers treat intuitions as evidence in Scenario A and Scenario B? Let us consider Scenario A first. The answers seem to be “Yes” and “No.” S treats intuitions as evidence, but P does not. Likewise, when considering Scenario B, the answers seem to be “No” and “Yes.” S does not treat intuitions as evidence, whereas P does. Thus, these two scenarios demonstrate that there might be a possible ambiguity in Centrality; does it matter what the author is doing, or should we primarily be concerned with what the reader of the article thinks upon reading it? If it is the latter, then strategy (b) would not be very informative.

Let us consider some ways of responding to the possible ambiguity in Centrality. First, it might be suggested that if we are analyzing S’ article to find out whether Centrality is true, what matters is what S is doing. What is going on in S’ article depends solely on what S is doing.

However, this assertion seems to disregard the dialectic aspect of philosophy. The point of writing philosophy articles is to present what one thinks is good evidence for a claim and convince the interlocutors that the available evidence favors that claim. This point might be made clearer by revising Scenario A, as follows:

Scenario A*: S thinks she is treating both A and I as evidence for C in the article she has written. In fact, this is what S is actually doing. However, P and the rest of S and P’s community do not think that S treats I as evidence for C. Rather, they think that S treats A as evidence for C. Thus, after having read S’ paper, P and the rest of S and P’s community accept C because of A, and only A.

In this case, it does not seem to matter much that S treats I as evidence. Nobody else considers I to be evidence for C. In the current context, the presence of I is idle and non-effective (Cappelen 2012, 117). S’ appeal to I as evidence does nothing to convince her interlocutors. Thus, if we want to make a general statement about the community described in A*, it might be accurate to say that in this community, I is not considered evidence for C.

If this assumption is right, it suggests that if one wants to know whether intuitions are treated as evidence for a certain claim, analyzing the article in which the claim is first defended and considering what the author is doing might not be the right approach. However, it does not automatically follow that we should instead look at what a reader of the article is doing, as scenario B* illustrates:
Scenario B*: S thinks she is treating A, and only A, as evidence for C. In fact, this is what S is actually doing. However, P is not convinced by A and believes that A does not amount to evidence for C. Yet, P thinks S treats I as evidence for C. Thus, after having read S’ article, P accepts C because of I, and only I. However, the rest of S and P’s community think that S treats A, and only A, as evidence for C. Thus, after having read S’ paper, the rest of S and P’s community accept C because of A, and only A.

If something like Scenario A* or B* represents the actual situation, and we studied what S does in A* or P’s response to S’ take on B*, then we would be studying anomalies. This suggests that what is important is not what the author or readers (in virtue of being the author or readers) of an article are doing, but what the majority is doing. (I will discuss issues concerning generalizing one’s findings about a sample to the population in Section 6.2.1.3. That is, I will discuss what, if any, reasons we have for believing that what a sample of philosophers is doing is what most philosophers in the population are doing.)

However, if we accept the assumption that finding out what the readers of an article are doing is crucial in an investigation of Centrality, then this might raise a further problem. It may be more complicated to provide relevant evidence than first assumed. For example, when a philosopher, X, builds an argument with a claim defended in the article of another philosopher, Y, as a premise, it might not be clear from reading X’s article what X treats as evidence for C. Given that C is defended in another article, X might not bother to explicate the evidence for C. Moreover, what Y is treating as evidence for the claim need not tell us anything about what X is treating as evidence for the claim—even when X believes that she is treating what Y is treating as evidence for the claim as evidence for the claim (as in Scenario B). Furthermore, given that what Y believes she is doing need not be what she is actually doing, what Y says that she does need not be good evidence of what X treats as evidence for C. Finally, as X might also have false second-order beliefs about what she does, what X says she is doing need not be what she is actually doing. In that case, neither strategy (a), (b) (here understood as analyzing the writings of an author), or (c) will necessarily provide us with the information we need to determine what X treats as evidence—and, for the same reasons, what the other readers of Y’s article are treating as evidence.

I will now describe how I solve the issues outlined above. I combine several strategies. One of the strategies I apply is strategy (b). For (b) to be a feasible strategy, I will need to deal with several issues. First, there are issues concerning making inferences about a population based on a study of a sample from that population. As mentioned, I will discuss that issue in Section
6.2.1.3. Second, attempts to apply strategy (b) have been troubled by the problem of underdetermination and the problem pertaining to demonstrating that claims about cases support theoretical accounts because of their intuitiveness. Articles 1 and 2 try to solve this problem by constructing a new framework for analyzing philosophy articles, which provides us with observations that allow us to determine whether (i)—claims about cases support theoretical accounts—is true. More precisely, the aim of the method is to provide observations that help us determine whether (i*) is true, that is, the claims about the cases provide some support for the theoretical accounts in virtue of having some credibility independently from its relation to the theoretical account. The difference between (i) and (i*) is that, according to (i), the support goes only from the claim about the case to the theoretical account, while (i*) is compatible with the support to some extent going in both directions. Moreover, the framework also offers diagnostics for deciding whether the theory-independent credibility of the claims about the cases comes from their intuitiveness.

The framework, then, helps us decide whether the intuition-deniers’ or friends of Centrality’s interpretation of the philosophical practice best fits the evidence (see Blau 2017, 262). Moreover, the framework provides clear guidelines for the analysis, thus preventing one reading too much into the text (see Blau 2017, 260).

The second strategy I apply is strategy (a). The main problem with strategy (a), however, is that philosophers’ second-order beliefs may be false. In Article 2, I argue that the upshot from this uncertainty is not that we should disregard what philosophers say about their method but rather that we must consider what they do and what they say together. If what they say is consistent with what they do, we have reason to believe that they do what they say they are doing—this relieves the last problem concerning strategy (b). Moreover, I consider the arguments offered by the intuition-deniers in defense of thinking that philosophers have false second-order beliefs about their method. I find these arguments wanting. Thus, I do not think we have reasons for believing that philosophers’ second-order beliefs are false. By developing a new method for applying strategy (b) and combining it with strategy (a), then, I will be able to overcome the challenges pertaining to making inferences about what the method of a philosopher is.\footnote{Although I might be able to make valid inferences about the author of the article’s method (i.e., what the author is actually doing), a final problem seems to be remaining: my approach does not seem to tell us anything about how the readers conceive of the articles analyzed, which is something the discussion above seems to suggest that we should take into account. By the end of Section 6.2.1.3, it will be clear why I do not think it is necessary to investigate the readers understanding of the articles.}
6.2.1.3 Case study research

The second level of inference concerns making inferences about a population based on studying a sample from that population. In this thesis, I conduct a small number of case studies. Therefore, a question that needs to be addressed is the following: To what extent am I warranted in generalizing my findings from those case studies to the population as a whole? This question is particularly relevant with regard to strategy (b), according to which we must consider what a philosopher does when she philosophizes.3 I proceed as follows: First, I briefly account for the reasons making case study a suitable approach. Second, I discuss the sampling technique applied by the intuition-deniers, which is important because my own approach is related to the strategy of the intuition-deniers. Third, I describe my own approach.

Commonly, inferences from a sample to an entire population are considered warranted when the sample is representative. Representative samples are often achieved by random sampling. However, to investigate (I), I conduct case studies. The reason why I conduct case studies is, first, the framework I develop requires assessing arguments in detail. Thus, my approach to analyzing the cases is quite intensive. The second reason is that even if less intensive techniques for analyzing cases were available, it would still be difficult to select a representative sample. For example, there is no straightforward way to define the population. The boundaries between political philosophy and neighboring disciplines are blurred, and it is not obvious whether scholars or publications should be treated as the units of analysis. Moreover, some persons and publications are more influential than others. This raises questions about whether and how persons or publications should be weighted to make our sample representative. These considerations speak in favor of doing case studies.

Let us consider Cappelen’s and Deutsch’s case studies of Centrality. Both pursue strategy (b). Cappelen (2012) claims that the thought experiments6 he studies are “paradigmatic” (96) and that they “are carefully chosen: they are exactly of the kind that proponents of Centrality mention (but don’t carefully study)” (130). For example, Thomson’s article on the trolley problem is selected because it is “one of the first and most cited papers on so-called ‘trolley [thought experiments],’” which have been “enormously influential and have triggered a great

---

3 It is relevant to strategy (a)—Consider what a philosopher says about her method—as well, but in this thesis, strategy (a) primarily amounts to providing arguments for why we can trust philosophers’ self-analyses and explaining how such reports can help us answer the question of whether Centrality is true. The thesis proceeds on the common place assumption that most philosophers have second-order beliefs in accordance with Centrality. Therefore, I will conduct the discussion of this issue in relation to strategy (b).

6 In the remainder of this section, I will refer to philosophical cases as thought experiments. This is to avoid confusion caused by using “case” to both refer case as in “case study” and case as in “philosophical case.”
deal of literature” and are “[b]y fairly wide consensus […] paradigms of important thought experiments in moral philosophy” (Cappelen 2012, 148).

Deutsch’s (2015b, 39-40) reasons for taking Gettier’s and Kripke’s arguments as case studies are as follows:

First, […] both Gettier’s argument against the JTB theory of knowledge and Kripke’s argument against the descriptivist theory of reference for proper names are targets of recent xphi [experimental philosophy] critiques. Second, although they concern issues in separate areas of philosophy (Gettier’s argument is epistemological while Kripke’s is semantic), both arguments share a structure, a structure they in turn share with an enormous range of other arguments in every branch of analytic philosophy. This shared argumentative structure allows methodological conclusions about Gettier’s and Kripke’s arguments to generalize to arguments in metaphysics, ethics, action theory, philosophy of mind, and so on. Third, it is quite common to hear misleading claims about the role that intuitions play in Gettier’s and Kripke’s arguments. If it can be shown how, in the cases of these arguments specifically, the claims are misleading, that ought to go a long way toward convincing philosophers that similar claims about other arguments in philosophy are likewise misleading.

Cappelen and Deutsch, then, make similar claims about the cases they study. The cases are supposed to be the cases that, if Centrality is true, would be the cases where we would be most likely to find appeals to intuitions about thought experiments. Moreover, the cases are supposed to be influential cases. Finally, the cases are supposed to be representative for philosophers’ use of thought experiments in general.

To further explicate the logic behind the case-selection techniques applied by the intuition-deniers, it might be helpful to consult the techniques for case-selection offered by John Gerring (2007) in his first edition of Case Study Research.⁷ Cappelen’s and Deutsch’s approaches seem to bear resemblance to a typical-case approach to case selection. According to Gerring (2007, 91), “The typical case exemplifies what is considered to be a typical set of values, given some general understanding of a phenomenon.” Moreover, “typicality may be understood according to mean, median, or mode on a particular dimension; there may be multiple dimensions […] and each may be differently weighted (some dimensions may be more important than others)” (Gerring 2007, 92). What is important for the present purpose is that according to the typical-

---

⁷ Although Gerring discusses his taxonomy with regard to case studies on causal relationships, the taxonomy is also helpful in the discussion of case-selection in descriptive case studies.
case approach, the general idea is that cases are selected based on knowledge about the characteristics of the population of cases.

In the study of Centrality, the typical-case approach would amount to defining the relevant population of cases, which would be thought experiments in contemporary analytical philosophy, identifying some features of the thought experiments, and studying those thought experiments that have the most common features. Cappelen and Deutsch appears to claim that they are doing something like this when selecting their cases. Cappelen claims that his cases are paradigmatic, and Deutsch claims that his cases have a structure that is typical for the population. (However, they do not elaborate on these claims.)

Perhaps a more interesting feature of their approach to case selection is that the cases they study are supposed to be those cases that friends of Centrality appeal to as evidence of Centrality being true. This feature alludes to another technique for choosing cases, which Gerring (2007, 115-122) describes as the crucial-case method for choosing cases. A case is deemed “crucial” “when it is most, or least, likely to fulfill a theoretical prediction” (Gerring 2007, 115). Thus, there are two kinds of crucial cases: “most-likely” cases and “least-likely” cases. Gerring (2007, 115) provides the following elaboration of these two kinds of crucial cases:

A “most-likely” case is one that, on all dimensions except the dimension of theoretical interest, is predicted to achieve a certain outcome, and yet does not. It is therefore used to disconfirm a theory. A “least-likely” case is one that, in all dimensions except the dimension of theoretical interest, is predicted not to achieve a certain outcome, and yet does so. It is therefore used to confirm a theory. (Italics in original)

The general logic of such cases is that “the crucial case offers a most-difficult test for an argument, and hence provides what is perhaps the strongest sort of evidence possible in a nonexperimental, single case setting” (Gerring 2007, 115).

The cases selected by Cappelen and Deutsch resemble most-likely cases. That is, they are supposed to be thought experiments Centrality predicts is most likely to be used to prompt intuitions that are treated as evidence. These cases, therefore, are suitable for falsifying

---

8 It is not uncommonly suggested that thought experiments in philosophy can have a variety of different functions (Brun 2018; Walsh 2011; Machery 2017, ch. 1; Brendel 2004). Thought experiments may, for example, be used to illustrate claims. If some thought experiments only function as tools for illustrating claims, this would not be the right kind of thought experiment to study if we want to determine whether Centrality is true. The thought experiments we want to study are those that philosophers make claims about that play an argumentative role. Perhaps, such thought experiments could be considered as exemplifying a “typical value” (that is, a typical characteristic) in the context of the study of Centrality and the investigation of philosophers’ use of thought experiments. (This, of course, presupposes that there are thought experiments that play an argumentative role.)
Groundwork for the Thesis

Centrality. However, Cappelen and Deutsch also make general positive claims about the method of contemporary analytical philosophers. These generalizations are warranted on the assumption that the cases they study are also typical cases. In this thesis, I grant that the cases studied by Cappelen and Deutsch can be categorized as both most-likely and typical cases.

Before I introduce my own approach, it is worth stressing that I investigate a narrower claim than Centrality. Centrality is a claim about contemporary analytical philosophy in general. However, in this thesis, I am primarily concerned with a particular branch of philosophy, that is, political philosophy. That said, the challenges concerning generalization still apply, but the population in my case is narrower than the population in the studies of Cappelen and Deutsch.

I apply several strategies to investigate whether political philosophers treat intuitions as evidence. First, I conduct a case study of two cases from political philosophy. I take it that these cases are typical cases, at the very least with regard to the literature on justice. The reason I think that these cases are typical is that, as I describe in Article 1, the cases I examine are representative of a common style of argument. These cases seem to share a structure with a large range of arguments in political philosophy more generally. Moreover, the cases I examine are considered highly influential. The cases I study, namely, the articles written by Ronald Dworkin and G. A. Cohen, sparked the now extensive literature on luck egalitarianism and have had a significant influence on the literature on egalitarianism in general. These are among the most cited articles in the literature on justice, cited 2599 and 3388 times, respectively, according to Google Scholar (17.05.2022). Moreover, Dworkin and Cohen are considered among the most influential political philosophers of our time. Thus, I find it reasonable to assume that many political philosophers have been inspired by Dworkin and Cohen’s way of philosophizing and that their way of philosophizing, in this sense, is typical.

Second, using the analytical framework I develop, I reexamine one of the cases studied by Cappelen: Thomson’s (1985) article on the trolley problem. This is one of very few cases from moral and political philosophy that have been studied with the purpose of investigating Centrality. There are several reasons why it is worth reexamining this case. First, this is not only an influential case in moral philosophy (Kamm 2015), in general, but it is also much cited

---

9 https://scholar.google.com/scholar?cites=1342622458272460420&as_sdt=2005&sciodt=0,5&hl=no [Accessed 17.05.2022].
10 https://scholar.google.com/scholar?cites=114194371093588283964&as_sdt=2005&sciodt=0,5&hl=no [Accessed 17.05.2022].
11 These cases are not crucial cases in the current context. To be that they would have to be least-likely cases; cases in which we would be least likely to find that intuitions are treated as evidence. However, such cases would not have been typical cases. Recall, according to Centrality the practice of treating intuitions as evidence is widespread. In choosing between typical and crucial cases, I here opt for the former.
in different strands of literature in political philosophy, in particular (see e.g. Frowe 2018; Otsuka 1998). Second, examining a case that is also considered by an intuition-denier allows for a direct comparison of my method and the method applied by the intuition-deniers. Based on this comparison, I argue that the intuition-deniers’ method has some shortcomings, which, in turn, suggests a more general critique of the intuition-deniers’ project. If this critique is sound, it is questionable whether the intuition-deniers have succeeded in providing any evidence suggesting that Centrality is false at all.

To summarize, my overall strategy is to combine strategies (a) and (b). To pursue (b), I develop an analytical framework, and to pursue (a), I look for an author’s explicit statements about how she argues in a particular article and examine her writings on her method. The reason why I am not applying (c) is that I think this would be superfluous, particularly because it is common ground in the current debate that Centrality is a widespread assumption, which suggests that most philosophers assume that their interlocutors treat intuitions as evidence.12 At the very least, studying some typical cases from political philosophy and arguing that intuitions are treated as evidence in these cases, questioning (and hopefully undermining) the findings and the methods of the intuition-deniers, and showing that we do not have reason to think that (political) philosophers have false second-order beliefs should be sufficient for pushing the burden of evidence onto the intuition-deniers.

6.2.1.4 *Descriptive claims and the problem of labelling*

According to Gerring (2012, 737), “[w]ith descriptive arguments in the social sciences it is often unclear what criteria might be invoked to disprove a proposition.” One particularly relevant problem that Gerring (2012, 738) notes is the problem he calls the “problem of labelling”:

---

12 A survey by J. R. Kuntz and J. R. C. Kuntz (2011, 649) seems to suggest that Centrality is not as widespread a belief as I assume. This is because only 50.9% of the surveyed philosophers (N = 282) “agreed that intuitions are useful to justification in philosophical methods” and “nearly 70% of [them] considered that intuitions are not essential to justification” (my italics). However, Wesley Buckwalter (2012) has noted a series of issues concerning the design of the survey, which make it difficult to interpret the data. Taking these issues into account, it is by no means clear that it follows that Centrality is not as widespread as I assume. Moreover, I have some additional concerns about the study. First, the sample is quite small. Hence, even disregarding the worries concerning the design of the study, we should be careful with generalizing the data generated by the study. Second, Kuntz and Kuntz do not report the results by area of specialization. Thus, we cannot know from the data they report how moral and political philosophers compare to those from other branches and philosophers in general. The philosophers surveyed come from 19 different branches of philosophy, and only 53 of them report ethics as their area of specialization. A final reason for not worrying too much about this survey is that that Centrality is a widespread belief is common ground for the debate in which I am engaged.
Any phenomenon of significance to social science is likely to call up multiple words, and multiple definitions for each of those words. Consequently, and because much is at stake in an author’s choice of terms and definitions, descriptive arguments are apt to remain essentially contested.

The point is that there are multiple definitions of the words used to describe a phenomenon of significance in the social sciences. To illustrate, consider the concept of “democracy.” Since there are different definitions of democracy, scientists may disagree on whether a particular state is democratic, despite agreeing on which facts obtain. What Gerring seems to think follows is that since the scientists do not disagree on the facts in this case, it is unclear how to falsify the claims made by the scientists—i.e., whether a particular state is (not) democratic. This is why “descriptive arguments are apt to remain essentially contested.”

If we follow the convention of the current debate about the role of intuitions in philosophy and explicate (I) by asking whether contemporary philosophers treat intuitions as evidence, the problem seems to be that there is a possibility that friends of Centrality and intuition-deniers “intuition” and “evidence” to mean different things. Thus, conducting case studies to discover several facts about how philosophers argue may not help at all in moving the debate forward. In this case, my dispute with the intuition-deniers would be verbal (Chalmers 2011). Our dispute over whether philosophers treat intuitions as evidence would, by hypothesis, arise wholly in virtue of the intuition-deniers and myself disagreeing about the meanings of “intuition” and “evidence” (Chalmers 2011, 522).

The question then becomes, “How should we respond to the possibility of there being a problem of labelling?” First, we need to determine whether a dispute is verbal or substantive. Second, if it turns out that the dispute is verbal, we should try to resolve it. Although Gerring thinks that descriptive arguments are apt to remain essentially contested due to verbal disputes, David Chalmers (2011), for example, has made several suggestions on how to resolve such disputes. Thus, trying to resolve verbal disputes need not be futile.

To determine whether a dispute is verbal, we may apply what Chalmers (2011, 526-527) calls “the method of elimination”:

To apply this method to a dispute over a sentence $S$ that is potentially verbal with respect to term $T$, one proceeds as follows. First: one bars the use of term $T$. Second: one tries to find a sentence $S'$ in the newly restricted vocabulary such that the parties disagree nonverbally over $S'$ and such that the disagreement over $S'$ is part of the dispute over $S$. Third: if there is such an $S'$, the dispute over $S$ is not wholly verbal, or at least there is a substantive dispute in the vicinity. If there is no such $S'$, then the dispute over $S$ is wholly
This would amount to barring the use of the terms “intuition” and “evidence” and reformulating the sentence “Philosophers treat intuitions as evidence” using the restricted vocabulary. The new sentence must be formulated in such a way that the intuition-deniers and friends of Centrality disagree nonverbally over it and that the disagreement over this newly formulated sentence is part of the dispute over the sentence “Philosophers treat intuitions as evidence.” If it is possible to formulate such a new sentence, the dispute over “Philosophers treat intuitions as evidence” is not wholly verbal. As the discussion in Section 4.3 suggests, it is possible to formulate such a sentence: “Philosophers treat verdicts about thought experiments as starting points for abductive arguments, in which the acceptance of a theoretical account figures as the conclusion, and where the starting points are granted some credibility independent of its relation to the theoretical account.” Thus, with regard to (I), the disagreement is substantive.

6.2.2 The Inquiry of Research Question (II)

The second part of the thesis is a philosophical inquiry into methodological issues: Does MRE presuppose the criterion of intuitive fit? Or, more precisely, do the arguments from experimental philosophy undermine MRE? In the following, I first describe my method before briefly discussing some methodological issues pertaining to that method.

6.2.2.1 Method

My approach is as follows: First, I outline the common beliefs among proponents of MRE about what characterizes the method, which they think is the proper method for moral theorizing. Second, I outline the set of beliefs that critics of MRE hold, that is, the beliefs that make up the challenge from experimental philosophy. This set of beliefs includes, for example, those pertaining to what properties intuitions have, whether those properties make intuitions suited for serving as data for moral theorizing, whether treating concrete intuitions as data is an essential feature of MRE, and whether MRE should be rejected. Moreover, the critics of MRE believe that if treating concrete intuitions as data is not an essential feature of MRE, then the method becomes vacuous.

Third, to determine whether the challenge from experimental philosophy undermines MRE, I consider how the beliefs in these sets square with each other and consider how plausible the beliefs within each set are. Crucially, one set of beliefs holds that MRE is the proper method for moral theorizing, while the other set of beliefs claims that MRE is not the proper method
for moral theorizing. According to the critics of MRE, the reason why MRE is not the proper method for moral theorizing is that it presupposes that concrete intuitions are suited for serving as data. Thus, an important question is whether it is true that the assertion that intuitions are suited for serving as data for moral theorizing is an essential feature of MRE.

Fourth, since I argue that according to some interpretations of MRE, treating concrete intuitions as data is not an essential feature of the method, I proceed to inquire into the claim that these interpretations of MRE are vacuous. According to this objection, construing MRE such that treating intuitions as data is not an essential feature of the method would result in a method without any distinctive features and that do not offer the necessary resources for adjudicating between methodological or substantive moral views. To inquire this objection, I clarify what critics mean by vacuous and then investigate whether this holds for the relevant interpretations of MRE.

6.2.2.2 Methodology

As described above, the starting point of my inquiry is two sets of beliefs that together make up an incoherent belief system. For example, the one set includes the belief that the assertion that concrete intuitions are suited for data is essential to MRE, while I will argue that the other does not. The central features of my method are to clarify claims and to investigate what one commits to by making a particular set of claims. To illustrate, I inquire whether the claim that “concrete intuitions are essential data for moral theorizing” is among the claims that make up MRE or can be deduced from those claims. Such an investigation relies on accepted rules of logic. Furthermore, it also relies on pointing out distinctions that may clarify the (set of) claims.

Similarly, when inquiring whether the set of beliefs that makes up MRE is vacuous, pointing out distinctions and applying rules of logic are important parts of my method.

By following this procedure, the coherence of the overall belief system may eventually be restored. Coherence may, for example, be restored by rejecting the claim that according to MRE, concrete intuitions are essential data for moral theorizing, rejecting the claim that MRE becomes vacuous if does not presuppose that concrete intuitions are essential data for moral theorizing, and therefore rejecting the claim that MRE should be rejected. In the case of (some versions) of the ideal of MRE, this is what I suggest we should do.

However, one might ask what reasons we have for accepting those claims and embracing MRE, rather than rejecting the claim that MRE is a plausible method, insisting either that an essential feature of MRE is that intuitions are treated as data or that the method is vacuous. This
alternative would also restore coherence. Thus, considerations of coherence alone do not seem to favor either option.

One way to respond to this issue is to stress that the aim of the current inquiry is to determine whether the challenge from experimental philosophy is sound. The challenge is only sound if all its premises are true (and the challenge is valid, which I grant it is). One of its premises is that according to MRE, concrete intuitions are essential data for moral theorizing. This claim can be tested by consulting what the proponents of MRE write about the method. If what they write about the method suggests that the premise is false, it would be odd to suggest that we should conclude that what the proponents of MRE write about the method is false. Thus, within the current dialectic, it makes sense to restore coherence by concluding that the challenge from experimental philosophy fails.

That said, things are a bit more complicated. This is because a practical implementation of a method may deviate from the ideal from which it is derived. In fact, Articles 1 and 2 back up the assertion that most philosophers treat intuitions as data. Thus, critics might suggest that we should consider what philosophers do instead of what they write about their method when considering the plausibility of the challenge. When the challenge is directed at what philosophers are doing, it seems to be sound (given that the rest of the premises are true as well). That is, when we look at how MRE is implemented, philosophers proceed as if it is true that concrete intuitions are essential data for moral theorizing. However, this remark does not change the fact that within the current dialectic, if, when consider what philosophers do, there is reason to believe that they do not treat intuitions as data, it would be reasonable to reject the challenge rather than revising our beliefs about what philosophers do to conform to the challenge.

(Note that this consideration also illustrates the importance of carefully pointing out relevant distinctions.)

Finally, there seems to be an even more fundamental problem lurking. I have claimed that rules of logic play an important role in the inquiry of (II). This raises an even more fundamental question: How are the rules of logic justified? In an often-quoted passage, Nelson Goodman (1983, 63-64) offers the following response:

How do we justify a deduction? Plainly, by showing that it conforms to the general rules of deductive inference. An argument that so conforms is justified or valid, even if its conclusion happens to be false. An argument that violates a rule is fallacious even if its conclusion happens to be true. To justify a deductive conclusion therefore requires no knowledge of the fact it pertains to. Moreover, when a deductive argument has been
shown to conform to the rules of logical inference, we usually consider it justified without going on to ask what justified the rules. […]

Yet, of course, the rules themselves must eventually be justified. […] Principles of deductive inference are justified by their conformity with accepted deductive practice. Their validity depends upon accordance with the particular deductive inferences we actually make and sanction. If a rule yields unacceptably inferences, we drop it as invalid. Justification of general rules thus derives from judgments rejecting or accepting particular deductive inferences.

This looks flagrantly circular. I have said that deductive inferences are justified by their conformity to valid general rules, and that general rules are justified by their conformity to valid inferences. But this circle is a virtuous one. A rule is amended if it yields an inference we are unwilling to accept; an inference is rejected if it violates a rule we are unwilling to amend. The process of justification is the delicate one of making mutual adjustments between rules and accepted inferences; and in the agreement thus achieved lies the only justification needed for either. (Italics in original)

This procedure resembles MRE. In fact, Rawls (1999, 18 n. 7) acknowledges that these remarks by Goodman are similar to his description of MRE in *A Theory of Justice*. This, then, suggests that MRE is the method for justifying logical rules (see also Føllesdal 1988, 117). Consequently, the worry is that the justification of my procedure depends on the soundness of method that itself is on trial.

One response could be to try to argue that logic and ethics are separate domains, and that logical intuitions are reliable. (However, note that some interpretations of MRE do not necessarily treat logic and ethics as separate domains (Føllesdal 1988, 117-118).) Another response, and the response that I find the most natural, is that we must treat some beliefs as justified to get our inquiries of the ground. In the case of the current inquiry, among the beliefs I treat as justified are the beliefs that make up the rules of logic. As John L. Pollock (1987, 4) argues in response to skeptical arguments questioning both our beliefs and our cognitive processes:

We cannot dispense with *both* the beliefs and cognitive processes, because then we would have nothing with which to begin again. As Otto Neurath […] put it in an often-quoted passage, “We are like sailors who must rebuild their ship upon the open sea.”[…] We must start with the beliefs and cognitive processes we have and repair them “from within” as best we can. (Italics in original, reference and note removed)
Presentation of the Articles

In this section, I present the articles of the thesis and explain how they relate to each other. I also make some suggestions for further research.

7.1 How the Articles Relate to Each Other

Articles 1 and 2 answer (I)—Do contemporary analytic political philosophers treat the criterion of intuitive fit as a theoretical desideratum? In Article 1, I develop an analytical framework for identifying whether a claim is treated as evidence because of its intuitiveness and apply this framework in investigating two articles from the field of political philosophy. The second article applies the framework in an inquiry of a third article and combines it with two further strategies for investigating whether philosophers treat intuitions as evidence. Thus, in addition to extending the analysis by examining a further article, Article 2 helps validate the framework developed in Article 1, thereby strengthening the analysis of (I).

Proceeding on the assumption that Articles 1 and 2 have established that contemporary analytic political philosophers treat the criterion of intuitive fit as a theoretical desideratum, Article 3 investigates (II)—Does MRE presuppose the criterion of intuitive fit? Critics of MRE seem to think that arguments from experimental philosophy situate proponents of MRE in a dilemma: either MRE presupposes the criterion of intuitive fit, which renders the method implausible, or it does not presuppose the criterion of intuitive fit, which leaves the method vacuous. Article 3 starts out by considering first horn of the dilemma, arguing that the arguments from experimental philosophy do not render (some interpretations of) the ideal of MRE implausible because they do not presuppose the criterion of intuitive fit. However, Article 3 also argues that MRE is not vacuous. Thus, (some) versions of MRE escape the apparent dilemma. That said, while Article 3 argues that (some interpretations of) the ideal of MRE do not presuppose the criterion of intuitive fit, it also argues that the challenge from experimental
philosophy can be construed as targeting a particular practical implementation of the ideal of MRE. That is, an implementation according to which the criterion of intuitive fit is treated as a theoretical desideratum. That such an implementation is quite common is backed up by the findings from Article 1 and 2 combined with the assertion that most moral and political philosophers apply MRE. When we consider (some versions of) the ideal of MRE, then, the answer to (II) is “No.” However, when we consider the current mainstream implementation of MRE, the answer is “Yes.” Since the current implementation of MRE treats the assumption that intuitions serve as data as fixed, the challenge from experimental philosophy undermines the current implementation of MRE. That is, the challenge pushes proponents of MRE to derive a new implementation from the ideal of the method according to which the criterion of intuitive fit is not treated as an essential theoretical desideratum.

To summarize, Articles 1 and 2 argue that political philosophers treat the criterion of intuitive fit as a theoretical desideratum. Departing from the conclusion of Articles 1 and 2, Article 3 argues that the criterion of intuitive fit does not play an essential role in the ideal of MRE. However, Article 3 also notes that the challenge from experimental philosophy has force against particular implementations of MRE according to which the criterion of intuitive fit is treated as essential. The findings of Article 1 and 2 (in combination with the assertion that most philosophers apply MRE) suggest that such implementations of MRE are quite common.

7.2 First Article

The article investigates whether Ronald Dworkin (1981a) and G. A. Cohen (1989) treat intuitions as evidence in the first articles in which they formulated and discussed versions of luck egalitarianism—“What is equality?” and “On the currency of egalitarian justice,” respectively. Due to Dworkin and Cohen have been highly influential in the field of political philosophy, and due to their arguments seeming to share a structure with a large range of arguments in political philosophy, I assume that these articles represent typical cases with regard to articles in political philosophy.

I developed an analytical framework for analyzing arguments to investigate whether Dworkin and Cohen treat intuitions as evidence. This framework aims to respond to the problems identified in Section 4. That is, the framework provides a strategy capable of providing positive evidence for Centrality. The framework involves two steps. The first step consists of finding out whether claims about cases are treated as evidence. If a claim about a

---

1 The article is published in The Journal of Political Philosophy (Conte 2022a).
case is treated as a starting point for an abductive argument in which the acceptance of a philosophical principle figures as the conclusion, then we have reason to believe that the claim about the case is treated as evidence for the philosophical principle. Thus, the framework provides diagnostics for identifying whether a claim is treated as a starting point for an abductive inference.

The second step consists of finding out whether the claim is treated as evidence because it is intuitive. The framework assumes that intuition-terminology being used to refer to a claim indicates that the claim is intuitive, and that the conjunction of the claim being treated as a starting point for an abductive argument and intuition-terminology being used to refer to the claim suggests that the claim is treated as evidence because it is intuitive.

The contribution of this article is manifold. First, it provides evidence that Dworkin and Cohen treat intuitions as evidence when articulating and discussing versions of luck egalitarianism. This, in turn, provides some reason to believe that intuitions are treated as evidence in political philosophy more generally. Second, it develops an analytical framework for investigating Centrality. This analytical framework, I argue, is capable of providing positive evidence for Centrality. More specifically, I argue that the Centrality-unfriendly explanation of the philosophical practice suggested by Cappelen and Deutsch fails to account for some features of philosophical method revealed by the findings of my analysis of Dworkin’s and Cohen’s articles. Third, the article also provides a tentative argument suggesting that Cappelen and Deutsch overlook features of the philosophical practice suggesting that Centrality is true. This argument is further developed in Article 2.

### 7.3 Second Article

In the second article, I investigate whether Thomson treats intuitions as evidence in her article “The Trolley Problem.” Not only is this a highly influential article, but it is also an article examined by Cappelen. Thus, examining this work allows for a direct comparison with Cappelen’s approach.

To investigate whether Thomson treats intuitions as evidence, I apply the framework developed in Article 1 in combination with several other strategies. In addition to examining Thomson’s arguments, I also look at what she explicitly claims she is doing in the particular article and consider her general writings about her method. Thus, the second paper draws on a wider collection of sources of evidence. This allows for a discussion of the validity of the

---

2 The article is published in *The Journal of Value Inquiry* (Conte 2022b).
different methods. If the three different methods deliver the same answer, this strongly suggests that the methods are valid.

Again, the contribution of this article is manifold. First, as already mentioned, it allows for validation of the analytical framework developed in Article 1. Second, by combining the three strategies, the article provides convincing evidence that Thomson treats intuitions as evidence. Third, the comparison with Cappelen’s analysis allows me to develop the argument suggested in Article 1; that Cappelen misinterprets some features of Thomson’s argumentation. I argue that what Cappelen believes are arguments for the claims about the cases are, in fact, abductive arguments in which the claims about the cases serve as the starting points. This argument implies a general objection to Cappelen’s and Deutsch’s case study approaches, since Deutsch’s approach is similar to that of Cappelen.

7.4 Third Article

In the third article, I consider whether the challenge from experimental philosophy is successful in undermining MRE. Critics of MRE suggest that the challenge only can be avoided by making MRE vacuous. Roughly, if MRE do not presupposes the criterion of intuitive fit, or more precisely, that concrete intuitions can be treated as data, it becomes a method without any distinctive features. This leaves proponents of MRE with the following dilemma: either MRE treats the implausible assertions that concrete intuitions are data for moral theorizing as essential or it is vacuous. Thus, in the third article, I consider whether there are any interpretations of MRE that escapes this dilemma.

To do so, the article draws upon several useful distinctions. First, it is common to distinguish between narrow and wide MRE. Very roughly, according to the narrow interpretation, we should build moral theories by bringing intuitions and principles into coherence. According to the wide interpretation, we should also include background theories among the elements that should be brought into coherence.

Second, the paper also distinguishes between different categories of background theories. That is, I distinguish between C-background theories—theories that primarily constrain the content of moral principles—and P-background theories—theories that primarily shape the reflective process. This distinction is important because, I argue, whether wide versions of MRE can withstand the challenge from experimental philosophy depends on what categories of background theories they include in the reflective process.

---

3 The paper is submitted to a scientific journal.
Third, noting that (wide) MRE is often claimed to be an unattainable ideal, the article distinguishes between the ideal of MRE and practical implementations of MRE—that is, what proponents of MRE do when they investigated substantive moral issues. For the purpose of Article 3, it is important to note that a particular practical implementation of wide MRE might be quite different from the ideal from which it is derived. Thus, we must distinguish between narrow and wide MRE, and, in particular with regard to wide interpretations of MRE, the ideal of MRE and the practical implementation of MRE. These distinctions help determining the implications of arguments from experimental philosophy.

I argue that the challenge from experimental philosophy successfully undermines narrow MRE. This is because narrow MRE presupposes P-background theories according to which concrete intuitions serve as data for moral theorizing. That intuitions serve as data is an essential feature of narrow MRE.

The challenge fails, however, when construed as targeting some interpretations of the ideal of wide MRE. In general, what wide interpretations of MRE that withstand the challenge have in common is that they include P-background theories in the reflective process. This means that they allow that reflective process itself can be revised in the course of an inquiry. For example, they allow for a critical stance toward concrete intuitions by including the findings from experimental philosophy studies as background theories. Moreover, according to the widest construal of MRE considered in the article, what we want our moral theories to represent and what is the proper data for moral theorizing, are themselves issues that must be decided in the reflective process. On this interpretation, to assume that MRE treat intuitions as essential would at best be premature.

Furthermore, expanding common responses to the charge of vacuousness, I argue that the ideal of MRE have several distinct features. For example, according to MRE, first-order issues cannot be decided independently of higher-order issues and vice versa. Moreover, claims from moral and non-moral domains are tested holistically.

Some versions of the ideal of wide MRE, then, escapes the dilemma. Narrow MRE, on the other hand, is undermined by the challenge from experimental philosophy. Since most proponents of MRE embrace the wide version of MRE, that narrow MRE is undermined seems like a not very significant victory for the challenge from experimental philosophy. However, I suggest while the challenge do not undermine the ideal of wide MRE, it can be construed as challenging an allegedly quite common practical implementation of it. That is, a practical implementation according to which the assumption that concrete intuitions serve as data for moral theorizing is treated as fixed.
7.5 Suggestions for Further Research

Several issues warrant further scrutiny. First, although this thesis assumes that the argument from experimental philosophy succeeds, experimental philosophy is still a quite recent and controversial field of research. Thus, there is still work to be done with regard to the argument from experimental philosophy. Such work is both empirical—concerning what properties intuitions have—and philosophical—concerning the epistemic premise of the argument from experimental philosophy. More specifically, it would be interesting if experimental philosophy studies investigate significant intuitions in political philosophy. Such a research project would involve, first, identifying significant intuitions in political philosophy (i.e., intuitions that have shaped the current debate), and, second, empirically investigating the properties of these intuitions. Moreover, Section 3.7 noted several issues concerning arguments from experimental philosophy that warrants further scrutiny.

Second, regarding the Centrality debate, the thesis only investigates a few articles. Thus, more case studies should be conducted.

Third, the thesis suggests that the reason why the role of intuitions is of particular theoretical interest is because intuitions are assumed to be a means to externally validate theories in moral and political philosophy. In the third article I argue that intuitions do not play an essential role in MRE. A question that warrants further examination is how to explicate versions of MRE not treating intuitions as data would look like and implement such versions of MRE.

Finally, the thesis suggests that while the challenge from experimental philosophy fails to undermine MRE as such, it nevertheless succeeds in challenging the current practice of many philosophers working on issues in political philosophy. However, this thesis does not investigate the implications of such (limited) success for first-order issues in political philosophy. Such an investigation could be a natural continuation of the line of research conducted in this thesis.
PART 2: ARTICLES
Are Intuitions Treated as Evidence? Cases from Political Philosophy*

8.1 Introduction

It is a common view that philosophers treat intuitions as evidence for philosophical theories. Following Herman Cappelen, we may refer to this view as centrality.¹ Advocates of centrality typically assume that claims about cases (henceforth, case verdicts) are treated as evidence for and against philosophical theories because of their intuitiveness. On the basis of their examination of prominent philosophers’ use of cases, however, critics of centrality, like Cappelen and Max Deutsch, claim that philosophers argue for case verdicts, and take that to suggest that advocates of centrality are mistaken in their assumption that the intuitiveness of the verdicts plays an epistemic role.² I’ll refer to these critics as “intuition deniers.”³ The intuition deniers’ rejection of centrality has launched a considerable debate over the nature of the philosophical method and the role of intuitions in philosophical argument.

¹ Cappelen 2012, 1.
² Ibid.; Deutsch 2015b; see also e.g. Ichikawa 2014; Molyneux 2014.
³ Following Nado 2016.

* This article is published in Journal of Political Philosophy (Conte 2022a). Earlier versions of this article were presented at the Department of Political Science, University of Oslo, the Nordic Network of Political Theory’s workshop (Roskilde 2018), the Norwegian national conference in political science (Trondheim 2019), and the ECPR general conference (Wroclaw 2019). The article benefitted greatly from discussions at these events and I am grateful to the participants for their comments. The article also benefitted from a stay at CAS as a GOODPOL-fellow. In particular, I would like to thank Kim Angell, Herman Cappelen, Andreas Christiansen, Keith Dowding, Göran Duus-Otterstrom, Jakob Elster, Fredrik Dybfest Hjorthen, Robert Huseby, Klemens Kappel, Sune Legaard, Erlend Langørgen, Sigurd Lindstad, Kasper Lippert-Rasmussen, Jonathon Moses, Tor Otterholt, and Dag Einar Thorsen, as well as Christian Barry and two anonymous referees for this journal, for valuable comments.

Copyright © 2010-2020 Knowledge Transfer Ltd. All rights reserved.
This article contributes to that debate by investigating cases from a genre of philosophy—political philosophy—that has not been the main focus so far. Within this branch (and within moral philosophy more generally), not only has it been claimed that intuitions are treated as evidence, but some claim that not doing so would leave the prospects of the discipline quite bleak. I respond to the challenge of providing positive evidence for centrality—a challenge posed by Cappelen—by developing an analytical framework. This framework will also help develop a common response to the intuition deniers’ arguments: that they mistake abductive arguments—in which case verdicts function as premises that support a theoretical account—for arguments that support case verdicts.

I examine cases from two influential articles in which some of the first attempts to formulate different versions of luck egalitarianism were made: Ronald Dworkin’s “What Is Equality? Part I” and Gerald A. Cohen’s “On the Currency of Egalitarian Justice.” I argue that a close examination of these articles indicates that their authors do treat intuitions as evidence. Both Dworkin and Cohen treat case verdicts as starting points for abductive arguments and use intuition-terminology ("intuition," "intuitively," and cognate terms) to express or to refer to these verdicts. I argue that the conjunction of these two observations constitutes positive evidence that they treat intuitions as evidence.

Cohen and Dworkin are viewed as being among the most influential thinkers in the literature on justice, and their discussion of expensive tastes is viewed as one of the most important debates in 1980s and 1990s political philosophy. Many regard them as two of the principal luck egalitarian thinkers. Due to their influence and importance, examining these articles is interesting in and of itself. Moreover, I think the current inquiry is of general interest. If centrality is false, we would not expect to find that intuitions figure as central evidence in influential contributions. Moreover, the issue Dworkin and Cohen raise and discuss in the articles I examine—the role of responsibility—pertains to distributive ideals in general. Finally, I think that Dworkin’s and Cohen’s articles are representative of a common style of argument. Their arguments seem to share a structure with a large range of arguments in political

---

5Tännström 2015, ch. 1; Tersman 2008, p. 404.
6Cappelen 2014a.
7Dworkin 1981a; Cohen 1989. While Dworkin (2003) does not view himself as a luck egalitarian and has explicitly rejected the label, many view him as one of the foremost luck egalitarian thinkers (see e.g. Knight 2009, 2013; Lippert-Rasmussen 2016; Stemplowska 2013).
8See e.g. Knight 2009, 2013; Lippert-Rasmussen 2016; Stemplowska 2013.
philosophy more generally. Thus, the articles seem representative of common methods of argumentation in political philosophy.\textsuperscript{10}

I proceed as follows. First, I provide some clarifying remarks on centrality and intuitions. Second, I provide a brief overview of the debate and present Cappelen’s challenge to provide positive evidence for centrality. Third, I develop an analytical framework in response to the challenge. Fourth, I apply the framework in an analysis of Dworkin’s and Cohen’s articles. Finally, I briefly consider some objections to the analysis and conclude.

\textbf{8.2 Centrality, Intuitions, and Cases}

My point of departure is Cappelen’s formulation of centrality:

\textit{Centrality (of Intuitions in Contemporary Philosophy).} Contemporary analytic philosophers rely on intuitions as evidence (or as a source of evidence) for philosophical theories.\textsuperscript{11}

Considering that centrality is an overly general claim to examine in a single article, I derive a narrower hypothesis from it, but first I will clarify how intuitions are understood in this study.

\textbf{8.2.1 A Note on Intuitions}

Intuitions are mental states formed non-inferentially.\textsuperscript{12} It is common to distinguish between two senses of intuition: intuitings and intuiteds.\textsuperscript{13} This distinction is important to the centrality debate, because the claim that philosophers treat intuitions as evidence is ambiguous. On the one hand, it can mean that philosophers treat intuitings as evidence: that is, they treat the occurrence of the mental state as evidence. On the other hand, it can mean that philosophers treat intuiteds as evidence: that is, the content of the mental state is treated as evidence.

In this article, the claim that intuitions are treated as evidence will be understood as follows. Intuiteds are treated as evidence, and they are treated as evidence because of the occurrence of intuitings.\textsuperscript{14} Since both Cappelen and Deutsch deny that intuitings serve as a source of evidence,

\textsuperscript{10}I am grateful to an anonymous referee and Christian Barry for pressing me to make the general interest of the article clearer.
\textsuperscript{11}Cappelen 2012, p. 3 (emphasis original).
\textsuperscript{12}E.g., Michael Huemer (2005, p. 102) views them as initial intellectual appearances. However, I will not make any strong commitments with regard to the nature of intuitions.
\textsuperscript{13}Lycan 1988, p. 211.
\textsuperscript{14}See Pust 2019.
they would say that, on this reading of the claim that intuitions are treated as evidence, the claim is false.\textsuperscript{15}

Finally, when inquiring whether intuitions are treated as evidence, I will follow Cappelen and Deutsch in focusing on the method of cases. That is, I will consider whether intuitions about cases are treated as evidence.\textsuperscript{16}

8.2.2 Hypothesis

The hypothesis that I investigate is a much narrower claim than centrality. First, I limit the study to a single genre of philosophy: political philosophy. Furthermore, I focus on the discussion of “What is equality?” and, more specifically, on the debate between Dworkin and Cohen about the first formulations of luck egalitarianism and the problem of expensive taste. This specification leaves us with the following hypothesis:

Intuitions served as central evidence when Dworkin and Cohen first articulated and debated luck egalitarianism.

I suggested above that Dworkin’s and Cohen’s articles represent a common style of argument. Thus, the narrowness of the hypothesis notwithstanding, the present inquiry is of general interest. However, due to the small number of articles examined, one should be cautious about generalizing the findings. To properly assess how much the current study’s findings can be generalized, follow-up studies would be needed, preferably with much larger samples.\textsuperscript{17}

8.3 The Intuitions Deniers and the Positive-Evidence Challenge

Based on examinations of philosophical practice, and the method of cases in particular, it has been argued that case verdicts are not supported by appeals to intuition.\textsuperscript{18} Instead, philosophers support such verdicts by arguing for them. Thus, the assumption is that a verdict being argued

\textsuperscript{15}Cappelen 2012, p. 3; Deutsch 2017, p. 424.
\textsuperscript{16}As an anonymous referee remarked, political philosophers sometimes express hostility toward the appeal to intuitions about toy cases, so it could be interesting to ask whether one can find the same kind of use of intuition as evidence in literature in which the author does not engage so straightforwardly in reasoning from such cases. The suggestion is interesting and worth pursuing, but due to the limited space available, I will not do so here.
\textsuperscript{17}Gerring 2017, ch. 10. Some large-scale studies have addressed themes in the centrality debate (see e.g. Ashton and Mizrahi 2018), but they do not address the specific challenge I outline below (Section 8.3).
for is a good indication that the verdict’s intuitiveness is not treated as evidence for the verdict. I refer to this position as the “argument view.”

A common reply to the argument view is that the idea that arguments support verdicts is compatible with the intuitiveness of the verdicts being treated as evidence. For example, the arguments’ role could be to shore up the intuition. Another common response is that the intuition deniers fail to see that some of the arguments they identify are actually abductive arguments with intuitions as starting points. Thus, support goes in the opposite direction to what the intuition deniers assume: that is, from case verdict to theoretical account. I will refer to the second response as the “explanation view.”

One problem with the first response is that it only establishes the possibility that intuitions are treated as evidence. However, more than that is needed to defend centrality, according to the intuition deniers. If a philosopher invokes arguments in defense of a case verdict, they claim, the verdict is supported regardless of whether its intuitiveness is treated as evidence. Thus, one needs positive evidence that the verdict’s intuitiveness is doing any work. I will refer to this challenge as “the positive-evidence challenge.”

The second response also falls short of the positive-evidence challenge, according to the intuition deniers. At best, those who have defended the explanation view would have showed that a given argument can also be interpreted as abductive. In that case, one should assume that the support goes in both directions and, thus, that the case verdict does not carry any independent evidential weight. Moreover, case verdicts are not necessarily treated as starting points for such inferences because they are intuitive.

8.4 Analytical Framework

As the discussion above demonstrated, friends of centrality need to confront the positive-evidence challenge to sustain their responses to the intuition deniers’ argument. It is useful to think of the positive-evidence challenge in terms of two hurdles. First, we need to demonstrate that case verdicts are treated as evidence. Second, we need to demonstrate that they are treated as evidence because they are intuitive.

---

8.4.1 First Hurdle
Although the explanation view might fall short of responding to the positive-evidence challenge on its own, the discussion above suggests a way of overcoming the first hurdle. We need diagnostics that help us demonstrate that case verdicts are treated as a starting point for abductive arguments because they carry some independent evidential weight. However, before I present such diagnostics, I will briefly explain how abduction is related to centrality and suggest an explanation as to why it is not always immediately clear whether an argument is abductive or not. This will help motivate the kind of framework that I develop in this article. It is a common notion that philosophy is based on an abductive methodology. This means that philosophers search for theoretical accounts that best explain the evidence. To see why it makes sense to look for abductive arguments when trying to figure out whether a case verdict is treated as evidence when we take this model as a starting point, we need a rough account of abductive arguments. It is common to assert that abductive arguments have something like the following structure:

\[ D \text{ is a collection of evidence (facts, observations, givens).} \]
\[ [A \text{ hypothesis}] \ H \text{ explains } D \text{ (would, if true, explain } E). \]
\[ \text{No other hypothesis can explain } D \text{ as well as } H \text{ does.} \]
\[ \text{Therefore, } H \text{ is probably true.} \]

Thus, “[t]he core idea is that a body of data provides evidence for a hypothesis that satisfactorily explains or accounts for that data (or at least it provides evidence if the hypothesis is better than explanatory alternatives).” Thus, given the popular idea that philosophy is an abductive enterprise and the fact that what is treated as evidence serves as starting points for abductive arguments, looking for whether case verdicts are treated as starting points for abductive arguments in which acceptance of a theoretical account figures as the conclusion is a suitable strategy for determining whether case verdicts are treated as evidence.

We now turn to the question of why it may not always be clear whether a case verdict is the starting point of an abductive argument (the explanation view) or is being argued for (the argument view). Stephen E. Toulmin’s distinctions between data, warrant, backing, and conclusion help to elucidate this. Conclusion is the claim “whose merits we are seeking to

---

25Ibid.
26Toulmin 2003, ch. 3.
establish.” Data are “the facts we appeal to as the foundation for [our conclusion]” and provide an answer to the question, “What have you got to go on?” Warrants are “general, hypothetical statements, which can act as bridges, and authorize the sort of step to which our particular argument commits us,” thereby providing an answer to the question, “How do you get [from the data to the conclusion]?”. Finally, backings are assurances standing behind our warrants “without which the warrants themselves would possess neither authority nor currency.” Data, warrant, and conclusion jointly are sufficient components of an argument, so not all arguments require backings.

In abductive arguments, case verdicts serve as data. The conclusion is that the theoretical account is plausible. What warrants the step from data to conclusion is the premise that the theoretical account is the best explanation of the case verdict. When case verdicts are argued for, the case verdict is the conclusion. The data are the facts of the case, and what warrants the step from data to conclusion is the theoretical account. Sometimes the warrant may be outlined in a way that makes it unclear whether the case verdict serves as data or conclusion, and considering that case verdicts and theories appear in both kinds of arguments, it might be difficult to determine which kind is present in a text. Moreover, a further complication, stressed by the intuition deniers, is that support might go in both directions. Thus, to overcome the first hurdle, we need a diagnostic that helps us determine whether case verdicts are treated as starting points for abductive arguments due to them having (some) independent evidential weight.

I propose the following set of diagnostics, which I call abduction. When a case verdict, $c$, is used as a starting point for an abductive argument, (1) not-$c$ is typically not a salient option in the context, (2) several theoretical accounts are under consideration, and (3) a theoretical account is considered more/less plausible if it can/cannot explain $c$, all else being equal, or (4) if more than one theoretical account can explain $c$, then verdicts ($c_2, \ldots, c_k$, where not-$c_2, \ldots, not-c_k$ are not salient options) about new cases are introduced as arbiters, all else being equal.

---

27Ibid., p. 90. Toulmin uses the term “claim” instead of “conclusion”; however, to avoid confusion with how I use “claim” elsewhere in this article, I use “conclusion” instead.
28Ibid., p. 90.
29Ibid., p. 91.
30Ibid., p. 90.
31Ibid., p. 96.
32See ibid., p. 98.
33The first two features are adapted from Cappelen’s (2012, pp. 122–4) discussion of the explanation view. Unfortunately, he does not pursue something like the strategy that I am suggesting. He discusses the explanation view in response to a suggestion that his diagnostics might not be able to discriminate between the explanation view and the argument view. Nothing that Cappelen says in his discussion of the objection undermines the strategy I have suggested for overcoming the first hurdle.
(1) suggests that the inquiry’s aim is not to determine how to judge the case in question, and (2) suggests that the inquiry’s aim is theory selection. The application of these features is pretty straightforward. To demonstrate that theoretical accounts are selected based on (among other things) their fit with case verdicts, we need (3) and (4). An indication of (3) being present is that an author treats a theoretical account’s fit or misfit with a case verdict as a reason for thinking that the theoretical account is more or less plausible—and perhaps even rejecting it. An author presenting the fit with a case verdict as a reason for choosing one theoretical account over another is an indication of (4). Thus, the presence of (1) and (2), in combination with (3) or (4), indicates that the case verdict supports the theoretical account.

8.4.2 Second Hurdle

If the abduction features are present, we have reason to believe that the case verdict is treated as evidence. However, we need another diagnostic for deciding whether the case verdict is treated as evidence because it is intuitive. I assert that if the case verdict is referred to or expressed using intuition-terminology, this is a reason to believe that it is treated as evidence because of its intuitiveness.

Some have suggested that intuition-terminology cannot be used for this purpose. For example, Cappelen argues that the ordinary use of intuition in English does not support centrality and that the philosophical—or “special technical”—use of “intuitive” and cognate terms is defective—roughly, philosophers do not use such terms in a distinct and clear way.34 In ordinary English, intuition-terminology is, for example, used by speakers to weaken their commitment to a claim or a conclusion. Regarding philosophers’ use of such terminology, Cappelen notes, for example, that there is no agreed-upon definition of “intuition.”35 However, although philosophers disagree on how to define intuition, they usually refer to the same phenomenon when they use the term.36

In this vein, John Bengson argues that “there is a discriminative use of ‘intuition’-terminology”37 that Cappelen fails to consider, which means

---

34Ibid., pt 1.
35Ibid., p. 52.
36Weinberg 2016; see also Deutsch 2015, pp. 30–2.
37Italics original.
in casual ordinary discourse, but which is not merely “special technical”—though it might be put to a theoretical purpose.\textsuperscript{38}

According to this understanding of intuition-terminology, “it denotes a conscious state or event—how one is appeared-to—that may play a substantial epistemic role.”\textsuperscript{39} Here are some examples of intuition-terminology from Bengson: “Intuition,” “intuitive(ly),” “counterintuitive,” “seem,” “appear,” “sound,” “look,” “see,” “perceive,” “strike,” “what we would say.”\textsuperscript{40}

Thus, following Bengson, I contend that such terminology is used to denote a certain conscious state; how one is appeared-to.\textsuperscript{41}

8.4.3 Taking Stock
To summarize, to determine whether case verdicts are treated as evidence because of their intuitiveness, we start by searching for the abduction features. If the abduction features are not present, the case verdicts are not treated as starting points for abductive arguments. If the features are present, which means we have reason to believe that case verdicts are treated as evidence, we then look for intuition-terminology. If present, we have reason to believe that they are treated as evidence because of their intuitiveness. Thus, my framework takes the form of a decision tree (see Figure 1).

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{analytical_framework.png}
\caption{Analytical framework}
\end{figure}

\textsuperscript{38}Bengson 2014, p. 559.
\textsuperscript{39}Ibid., p. 562.
\textsuperscript{40}Ibid., p. 561.
\textsuperscript{41}See also Devitt 2015.
8.5 The Method of Cases and Expensive Tastes

In their articles, Dworkin and Cohen articulate accounts of the political ideal of equality—equality of resources (ER) and equal access to advantage (EAA), respectively—that are commonly labeled luck egalitarian. ER holds that resources should be distributed equally, whereas EAA holds that access to advantage should be distributed equally. These accounts are developed in response to a third account of equality, equality of welfare (EW), which holds that the proper currency of egalitarian theories is welfare. What Dworkin and Cohen find unsatisfactory about EW is that it does not take into account an agent’s responsibility. This is something they try to remedy when formulating their own theories. Their discussions of expensive tastes are critical to their argument and I will focus on them here.

I start by outlining Dworkin’s and Cohen’s discussion of expensive tastes, particularly the Louis case, in some detail, italicizing parts of the quotes that are important for the discussion to follow. This is necessary to provide the reader with the relevant passages for interpreting the arguments. I then illustrate the difficulty in determining whether case verdicts are argued for or treated as starting points for abductive arguments, which stress the need for the analytical framework I developed above. I then use that framework to analyze the articles and test my hypothesis.

8.5.1 Dworkin

Dworkin aims to debunk EW and vindicate ER.\(^{42}\) One of the objections that Dworkin has against EW is the problem of expensive tastes:

\[(Q1) \text{[EW]} \text{ seems to recommend that those with champagne tastes, who need more income simply to achieve the same level of welfare as those with less expensive tastes, should have more income on that account. But this seems counter-intuitive, and I said that someone generally attracted to the ideal would nevertheless wish to limit or qualify it so that his theory did not have that consequence.}^{43}\]

So, for EW to be an attractive ideal, it must be qualified so that it avoids this problem without conceding “the final irrelevance of the principle.”\(^{44}\) Whether this is possible is what Dworkin wants to determine. After having raised this question, he presents the Louis case:

\(^{42}\) ER is elaborated further in Dworkin 1981b.
\(^{43}\) Dworkin 1981a, p. 228.
\(^{44}\) Ibid., p. 229.
(Q2) Imagine that a particular society has managed to achieve equality of welfare in some chosen conception of that ideal. Suppose also that it has achieved this through a distribution that in fact (perhaps just by coincidence) gives everyone equal wealth. Now suppose that [Louis] sets out deliberately to cultivate some taste or ambition he does not now have, but which will be expensive in the sense that once it has been cultivated he will not have as much welfare on the chosen conception as he had before unless he acquires more wealth. … Can Louis be denied extra wealth, taken from those who acquire less expensive tastes (or simply keep those they already have), without contradicting the ideal of [EW] that his community has embraced?45

Before answering this question, Dworkin adds another detail to the case:

I want to suppose that Louis is not only acting deliberately rather than inadvertently, but is also acting on the basis of the kind of judgment I said people often make when they form and change their preferences. He is trying to make his life a better life in some way.46

He then presents his judgment:

(Q3) This does not make his claim for extra resources any more appealing or less counter-intuitive, I think. On the contrary, the fact that he is acting so deliberately in his own interests seems to make his claim, if anything, less appealing than the claim of someone who tries an expensive experience on a whim, for the pleasure of the moment, and then finds that he is hooked.47

To block the counterintuitive implications, EW must be qualified by some principle. Dworkin considers several principles. First, he considers the principle of utility. However,

(Q4) the principle of utility does not explain what needs explaining here. It can at best explain why compensating those who develop expensive tastes is inefficient. It cannot explain why the ideal of equality does not recommend doing so.48

He then goes on to argue that “it is far from plain that the utilitarian principle, by itself, can even provide an explanation of what it does purport to explain”49 and concludes that:

46Ibid., p. 230.
47Ibid. (my italics).
48Ibid., p. 235 (my italics).
49Ibid.
(Q5) the supposed utilitarian justification of our intuitive conviction, that equality does not require that those who deliberately cultivate expensive tastes have equal welfare after they have done so, fails on two grounds. We still lack a justification for that conviction.\(^{50}\)

He then considers whether one can “argue in the following way”:\(^{51}\)

(Q6) Louis knows, or at least ought to know, that if he cultivates some expensive taste in a society dedicated to equality of enjoyment, for example, and is compensated, then that will decrease the enjoyment available for others. If, knowing this, he chooses the more expensive life then he does not deserve compensation. He is no longer a member of the company of those who deserve equal enjoyment in their lives.\(^{52}\)

Dworkin “find[s] this argument both powerful and appealing.”\(^{53}\) Moreover, he thinks that “[i]t is also an important argument for the following reason,” which is that the objection to giving Louis extra resources “is most naturally put in this way”:

(Q7) Louis should be free (at least within the limits allowed by a defensible form of paternalism) to make the best sort of life he can with his fair share of social resources. But he should not be free to trespass on the fair shares of others, because that would be unfair to them.\(^{54}\)

So, Dworkin finds the argument plausible because it may be grounded in the idea of fair shares. However, fair shares cannot be defined in terms of welfare, because then one would have no grounds for refusing Louis’s request for extra resources.\(^{55}\) Moreover, according to Dworkin,

[i]f the idea of fair shares is to do any work, then it must appeal to some independent account of fairness in distribution, and any independent account contradicts the conception to which it is attached … because it occupies all the space that conception claims for itself.\(^{56}\)

---

\(^{50}\)Ibid., p. 237(my italics).

\(^{51}\)My italics.

\(^{52}\)Ibid., p. 237 (italics original).

\(^{53}\)Ibid., p. 238.

\(^{54}\)Ibid.

\(^{55}\)Ibid., p. 234.

\(^{56}\)Ibid., p. 239.
Thus, a proponent of EW cannot appeal to the idea of fair shares, because defining fair shares in terms of something other than welfare would lead to a contradiction. The independent account of fair shares that Dworkin thinks is the most plausible is some conception of ER.\(^57\) Thus, if one rejects EW and accepts ER instead, then one can provide an explanation or an interpretation of the argument that Louis does not deserve extra resources.

Dworkin then considers whether

ingenious could produce some explanation or interpretation of the argument in question—that Louis does not deserve more resources just because he has chosen a more expensive life—which does not use this idea of fair shares or any similar ideas.\(^58\)

However, he suspects that any such account “would … fall before the following further example” and introduces the Jude case.\(^59\) Like Louis, Jude lives in a society dedicated to equality of enjoyment and decides to develop new and more expensive tastes because he wants to make his life better in some way. However, unlike Louis, Jude has far less money than the other members of his society because his original wants were so simple and inexpensively satisfied. Jude is now asking for more funds to satisfy his new wants through a redistributive scheme that would still leave him with fewer funds than anyone else. Dworkin judges the case as follows:

\[Q8\] Do we now have any grounds for saying that he is undeserving of the increase, when we know that if it is denied he will have both less funds and less enjoyment than anyone else? I doubt anyone will want to say this.\(^60\)

He thinks this demonstrates that “we cannot say that the reason Louis is undeserving of an increase is simply that the taste he has cultivated is expensive” because “Jude’s new taste may be just as expensive.”\(^61\) To express the force of the difference between the Jude and Louis cases, we need “the idea of fair shares (in this particular case the idea of an equal share of resources).”\(^62\)

---

\(^57\)Ibid.
\(^58\)Ibid.
\(^59\)Ibid.
\(^60\)Ibid. (my italics).
\(^61\)Ibid.
\(^62\)Ibid.
8.5.2 Cohen

In his article, Cohen challenges EW, ER, and equality of opportunity of welfare (EOW), a view developed by Richard Arneson, according to which opportunities for welfare are what should be distributed equally. Cohen offers EAA, which he views as a superior articulation of the ideal of equality.

The first part of Cohen’s discussion of Dworkin’s article is concerned with a case about a doubly unfortunate person. He believes “that egalitarians will be moved to compensate him for both of his misfortunes, but the fact that the first calls for egalitarian compensation challenges [EW] and the fact that the second does challenges [ER].” The discussion of this case leads Cohen to “affirm [EAA]” and he concludes with the following claim:

(Q9) (I warned at the outset that my positive proposal would be crude. One thing that makes it so and makes me wish that it will be superseded is the unlovely heterogeneity of the components of the vector of advantage. One hopes that there is a currency more fundamental than either resources or welfare in which the various egalitarian responses which motivated my proposal can be expressed. But I certainly have not discovered it, so, at least for now, I stay with the appearances, which contradict welfare, resources, and opportunity for welfare readings of the egalitarian demand, and which point, in the first instance, to the theory [or semitheory: it is perhaps too close to the intuitive phenomena to merit the name “theory”] I have affirmed.)

Cohen then examines how ER and EAA deal with expensive tastes. In this part of the discussion, he considers the Louis case and briefly outlines the case’s central features before writing that:

(Q10) while Dworkin and I both refuse Louis’s request for a special allowance, we ground our refusals differently. Dworkin says: sorry, Louis, we egalitarians do not finance expensive tastes; whereas I say: sorry Louis, we egalitarians do not finance expensive tastes which people choose to develop.

Cohen agrees with Dworkin’s judgment about the Louis case, but “grounds” it differently. In fact, he thinks that Dworkin “rejects the most obvious reason” for denying Louis extra

---

63 Arneson 1989.
64 Cohen 1989, p. 917.
65 Ibid., p. 921 (brackets original, my italics).
66 Ibid., p. 923 (my italics).
resources: “that he ‘sets out deliberately to cultivate’” his expensive tastes. Moreover, Cohen does not agree with Dworkin that “only [ER] can explain why Louis’s expensive tastes should not be indulged by egalitarians,” noting that one would be able to do so if one defined fair shares in terms of EOW:

(Q11) we could say that shares are fair when they equalize welfare opportunities. It is therefore false, and it scarcely takes ingenuity to show it, that only if we move toward [ER], toward fair shares in Dworkin’s special sense, can we explain egalitarianism’s lack of sympathy for Louis.

Thus, “the proposals that equality means [EOW] … glides by the Louis counter-example” and, a fortiori, so does EAA.

Cohen also discusses the Jude case, which he thinks shows that EAA is a better interpretation of the ideal of equality than both ER and EOW:

(Q12) Pace [EOW], I see no manifest injustice in Jude’s getting the funds he needs to travel to Spain. He then still has fewer resources than others, and only the same welfare, so [EAA] cannot say, on that basis, that he is overpaid. But, pace [ER], it seems not unreasonable to expect Jude to accept some deduction from the normal resource stipend because of his fortunate high ability to get welfare out of resources. Unlike either Dworkin’s theory or Arneson’s, mine explains why both gross underresourcing and gross “underwelfaring” (despite, respectively, a decent welfare level and a decent resource bundle) look wrong.

Note that Cohen makes two verdicts about the Jude case. He “sees” no injustice in Jude’s getting funds to travel to Spain, and it does not “seem” unreasonable to him to expect Jude to accept some deduction from the normal resource stipend.

67Ibid. (my italics).
68Ibid. (my italics).
69Ibid., p. 924 (my italics).
70Ibid. (note omitted).
71Ibid., p. 917.
72Ibid., p. 925 (my italics).
8.5.3 Searching for Arguments

At first glance it is not clear whether case verdicts are argued for or treated as starting points for abductive arguments in these articles. I will use Dworkin’s discussion to illustrate this point.\(^{73}\)

Here is one interpretation of Dworkin’s argument based on the quotes above:

\(P1\). Louis chooses to develop expensive tastes and “knows, or at least ought to know, that if he cultivates some expensive taste in a society dedicated to equality of enjoyment, for example, and is compensated, then that will decrease the enjoyment available for others” (Q6) (Data).

\(P2\). If one knows, or ought to know, that cultivating some expensive taste and being compensated decreases the enjoyment of others and still cultivates such tastes, then one does not deserve compensation (Q6) (Warrant).

\(P3\). One “should not be free to trespass on the fair shares of others, because that would be unfair to them” (Q7) (Back­ing).

\(C\). Louis does not deserve compensation (Conclusion).

Some observations seem to support this interpretation. Dworkin explicitly refers to Q6 (from which P1 and P2 are derived) as an “argument.” He claims that he “find[s] this argument both powerful and appealing,” and that it is “an important argument.”\(^{74}\) Moreover, Dworkin claims that the argument (P1, P2; therefore C) is important “for the following reason,”\(^{75}\) and then presents Q7, from which P3 is derived. This could suggest that P3 serves as an assurance standing behind the warrant.\(^{76}\)

However, another interpretation is possible. Dworkin may be treating the case verdict as a starting point for an abductive argument. First, Dworkin states the Louis verdict “needs explaining” (Q4). Second, Q6 and Q7 can be interpreted as Dworkin offering the idea of fair shares as an explanation of the case verdict. Third, from Q1 and Q3, it is clear that EW is not consistent with the Louis verdict, and from Q4, Q5, and Dworkin’s discussion of fair shares (outlined in the paragraph following Q7), it is clear that Dworkin thinks EW cannot be

\(^{73}\)I believe that the same point applies to Cohen’s discussion.
\(^{74}\)Dworkin 1981a, p. 238 (my italics).
\(^{75}\)Ibid. (my italics.)
\(^{76}\)It should be noted that Dworkin thinks that, as mentioned in Section 8.5.1, if the idea of fair shares is to “do any work” (which presumably includes serving as backing), fair shares should be defined in terms of equality of resources.
compromised in such a way that it is consistent with the Louis verdict. The argument would then be as follows:

\[ R1. \] Louis does not deserve extra resources (Data).
\[ R2. \] The idea of fair shares (understood in terms of ER) can explain that Louis does not deserve extra resources (Warrant).
\[ R3. \] EW cannot explain R1 (Warrant).
\[ D. \] Therefore, the idea of fair shares is the most attractive conception of equality (Conclusion).

If we compare the first interpretation with the second, we see that the case verdict is the conclusion in the former, but the starting point of the latter. While the idea of fair shares serves as backing in the first interpretation, it figures in the conclusion of the latter. The first interpretation suggests the argument view, while the second suggests the explanation view. At first glance, there are no conclusive reasons for favoring one interpretation over the other. To determine whether the case verdict is treated as supporting the theoretical accounts due to their independent evidential weight, we need to apply the framework developed above.

### 8.5.4 Abduction

To reiterate, the four features of abduction are: (1) not-\( c \) is not a salient option in the context, (2) several theoretical accounts are under consideration, and (3) a theoretical account is considered more/less plausible if it can/cannot explain \( c \), all else being equal, or (4) if more than one theoretical account can explain \( c \), then verdicts about new cases are introduced as arbiters, all else being equal. I asserted that the presence of (1) and (2) in combination with (3) or (4) indicate that \( c \) is treated as evidence. Let us consider Dworkin’s and Cohen’s arguments.

First, compensating Louis for expensive tastes is not a salient option in either article. At no point does Dworkin consider arguments for compensating Louis. Of course, he notes that EW demands that we compensate Louis, but rather than viewing this as a reason for compensating Louis, he suggests that this implication of the theory is problematic and counts against the theory (see (3)). Likewise, Cohen does not consider any arguments for transferring resources to Louis and thinks that Dworkin’s argument successfully defeats EW.

Second, several theoretical accounts are under discussion in both Dworkin’s and Cohen’s articles. When inquiring how to qualify EW so that it avoids the counterintuitive implications, Dworkin considers both the principle of utility (Q4, Q5) and the idea of fair shares understood
in terms of ER (Q6, Q7). Cohen, in addition to considering EW and ER, considers EOW and EAA (Q9).

Third, both Dworkin and Cohen seem to think that whether a theoretical account can explain the Louis verdict has direct bearing on its plausibility. For example, Dworkin finds EW implausible because of its implications for the Louis case (Q4). Cohen thinks the fact that EOW and EAA are consistent with the Louis verdict means that ER is not the only plausible account (Q11).

Fourth, although the Louis verdict is not introduced as an arbiter, but rather as a challenge to EW, Dworkin and Cohen treat other cases as arbiters. For example, Dworkin treats the Jude case as an arbiter when he considers whether the Louis verdict can be explained by an account that does not use the idea of fair shares defined in terms of equality of resources. As outlined in Section 8.5.1, Dworkin doubts that such an account can explain the verdicts in both cases. While other accounts might be consistent with the Louis verdict, one needs to define fair shares in terms of resources to account for the Jude case. And, importantly, saying that Jude is undeserving of the increase in funding is not a salient option. Thus, Dworkin thinks that ER can explain the difference between the two cases and that this is a reason for preferring it over theories that cannot do that. In other words, the Jude case is used as an arbiter between ER and other principles that might account for the Louis verdict.

Now note how Cohen proceeds. It is clear from Q9 that Cohen thinks that the fact that EAA “stays” with the appearances that contradict the other theories is a reason for favoring EAA over EW, ER, and EOW. “Appearances” in this quote refer to case verdicts, among them the case verdict about the doubly unfortunate person. Not compensating the unfortunate person is not a salient option. Moreover, Cohen thinks that EAA provides a better explanation of the verdicts about the Jude case (for which there are no salient options) than EOW and ER, and treats this as a reason for favoring EAA (Q12). This suggests that these case verdicts are treated as arbiters.

Thus, the presence of the abduction features provides us with reason to think that the case verdicts are treated as starting points for abductive arguments and hence as evidence. This suggests that Dworkin and Cohen’s arguments have the structures outlined in Table 1.
Dworkin

$D_1$ is a set of case verdicts (Louis, Jude, etc.);

ER can explain $D_1$;

EW cannot explain $D_1$;

Therefore, ER is the most attractive conception of equality.

Cohen

$D_2$ is a set of case verdicts (Louis, Jude, the unfortunate person, etc.);

EAA can explain $D_2$;

EW, EOW, and ER cannot explain $D_2$;

Therefore, EAA is the most attractive conception of equality.

Table 1. Dworkin’s and Cohen’s arguments

8.5.5 Intuition-Terminology

Q3 is the central quote for assessing whether Dworkin uses intuition-terminology when referring to the Louis verdict, as this quote renders his judgment of the Louis case.\(^{77}\) In this quote, Dworkin uses terms like “counter-intuitive” and “seems” when expressing the verdict (Q3). This suggests that it appears to Dworkin that Louis does not deserve extra resources. Moreover, Dworkin also uses intuition-terminology—“anyone will want to say”—when referring to the judgment about the Jude case (Q8).

Cohen also uses intuition-terminology. In Q9, he uses intuition-terminology to refer to the case verdicts, such as the verdict about the unfortunate person. The case verdicts are referred to as “appearances” and “intuitive phenomena.”\(^{78}\) He also uses terms such as “seems” and “look” when accounting for the verdicts about the Jude case (Q12).\(^{79}\)

\(^{77}\)There are instances of intuition-terminology in Q1 and Q5, but I presume that this terminology could be referring to a more general conviction: that we should not compensate expensive tastes in general.

\(^{78}\)See also Cohen 1989, p. 918.

\(^{79}\)Although Cohen does not explicitly refer to the Louis verdict using intuition-terminology, there are some reasons for believing that he treats it as a starting point for an abductive argument because it is intuitive. First, Cohen often refers to case verdicts using intuition-terminology, which could suggest that he generally finds case verdicts intuitive. Second, considering that Cohen is responding to Dworkin’s article, and agrees with Dworkin’s verdict about the Louis case, it seems reasonable to believe that he also accepts whatever Dworkin accepts as evidence for the verdict. In any case, although some might be unpersuaded by these remarks, Cohen clearly refers to both the verdict about the unfortunate person and the Jude case using intuition-terminology, so there should be no doubt that these verdicts pass the test provided by the analytical framework.
8.5.6 Taking Stock

The presence of the abduction features gives us reason to believe that the case verdicts are being treated in these arguments as having independent evidential weight, and the use of intuition-terminology gives us reason to believe that they are so treated because they are intuitive.

8.6 Objections and Replies

In this section, I consider some objections to the diagnostics that I have applied—that is, abduction and intuition-terminology.

8.6.1 Abduction

Some might not be convinced that the presence of the abduction features favors the explanation view. They might want to object that (1) not-\(c\) is typically not a salient option in the context and (2) the presence of several theoretical accounts could suggest that Dworkin and Cohen are providing several arguments for the case verdicts. Someone presenting this objection may appeal to Q5. After having concluded that the principle of utility cannot explain the Louis verdict, with no other argument or explanation for the verdict yet put forth, Dworkin claims that “We still lack a justification for that conviction” (Q5). The objectors could claim this suggests that the Louis verdict is unsupported unless we can provide an argument to support it, and that Dworkin cannot be treating the case verdict as evidence.

Consider first the statement “lack a justification.” In replying, one may point out that if the verdict lacks any credibility whatsoever, it should pose no problem for EW. However, it is clear that Dworkin thinks that EW is in trouble.\(^\text{80}\)

Second, it is true that the argument view could be consistent with features (1) and (2). However, the presence of feature (3)—a theoretical account is considered more/less plausible if it can/cannot explain \(c\), all else being equal—contradicts the argument view, since it suggests that the direction of support goes from the case verdict to the theoretical account. If the direction of support went only from the theoretical account to the case verdict, Dworkin would not reject

\(^{80}\)Moreover, it is commonly assumed that the method of reflective equilibrium, often claimed to be the most common method in moral and political philosophy (List and Valentini 2016, p. 542; McMahan 2013, p. 110), is committed to treating coherence as a necessary condition for justification. This is compatible with the idea “that a moral intuition carries at least some initial evidential weight”; Lycan 2019, pp. 108–9. Thus, the “lack of justification” statement need not suggest that intuitions are not treated as evidence, but rather that another condition that bears on the justification of a belief system—e.g., coherence with theoretical accounts—is lacking. Thus, I need not deny that case verdicts can receive support from theoretical accounts. However, my aim has been to demonstrate that some case verdicts are treated as having some evidential weight (independent of whether they cohere with a theoretical account), which I contend the abduction features—particularly (3) and (4)—demonstrate (see below).
ARE INTUITIONS TREATED AS EVIDENCE?

theoretical accounts that are inconsistent with the case verdict. Rather, we would expect the case verdict to be derived from the theoretical account. And if the relation of support between the case verdict and the theoretical account was merely mutual, lack of fit alone would not be a reason for rejecting either. However, as argued above, Dworkin treats the lack of fit between EW and the Louis verdict as a reason for taking EW to be implausible.

Finally, there is a further problem with the argument view. Both Dworkin and Cohen try to answer the question, “What is equality?” That is, their aim is to formulate theoretical accounts. Thus, if the Louis case and the other cases are used to vindicate these accounts, the accounts cannot themselves figure as premises in arguments for the case verdicts if the argument is our only reason for believing the case verdicts. If this were the case, the arguments would be viciously circular. 81

8.6.2 Intuition-Terminology

It could be objected that I have not sufficiently justified my reliance on intuition-terminology. Had I not relied on it, I could not have excluded the possibility that the case verdicts I have discussed are treated as starting points for abductive arguments because they are in the common ground—that is, they are claims that the participants in the debate agree on and are committed to—and they are in the common ground for reasons other than their being intuitive. If I cannot exclude this possibility, the argument goes, I should assert that the verdicts are in the common ground for some extra-intuitive reason, because intuition is “a controversial and elusive category,” whereas “there is no controversy over” claims “that happen … to be in the common ground.” 82

The argument presupposes that there are extra-intuitive sources of common ground claims. I will consider some possible extra-intuitive reasons why claims may be in the common ground. 83 First, claims could be in the common ground because they are facts. 84 However, as Deutsch notes, this only relocates the problem, as we now may ask, “How do we know that it is a fact?” In the moral domain, philosophers disagree on whether moral facts exist. Although moral realism—the view that moral facts exist—seems to be quite popular, 85 some realists argue

81 Of which Deutsch (2017, p. 440, n. 11) is aware.
82 Cappelen 2012, pp. 154–5.
83 See also Nado (2016, pp.796-7) for an argument suggesting that Cappelen fails to articulate any extra-intuitive sources for common ground claims.
84 Deutsch 2015, p. 45.
85 Bourget and Chalmers 2014; Yaden and Anderson 2021.
that moral facts are accessed via intuition. Thus, it is not clear that facts function as an extra-intuitive source of common ground claims in the moral domain.

Alternatively, claims could be in the common ground because very good arguments have been previously presented for them. To this, my response is that the cases I have discussed appear for the first time in the articles that I am analyzing. Thus, arguments have not previously been presented for these specific claims. A critic may object to this response by suggesting that while the cases are novel, the arguments need not be. The cases may be used to illustrate widely accepted arguments. However, this attempt to push back the objection fails. Dworkin and Cohen do not agree on why Louis should be denied extra resources (Q10).

Perhaps tacit knowledge could be the source of case verdicts that are in the common ground. According to Cappelen, we have pre-existing tacit knowledge that we may apply to novel cases. Moreover, he argues that judgments produced by tacit knowledge are not intuitions: “The reflections that generate [tacit] knowledge … are slow and under control,” whereas intuitive thinking is spontaneous and uncontrolled. Furthermore, he claims that he “know[s] of no usage of ‘intuitive’ that would make appeals to tacit knowledge of the kind described above intuitive.”

However, it is not clear that this source of common ground claims is distinguishable from intuition. Several accounts of moral intuition hold that moral intuitions are based on knowledge of the kind Cappelen describes. Knowledge underlying moral intuition, according to these accounts, can be acquired consciously, while eliciting spontaneous mental states: “Over time, previously undertaken episodes of conscious reasoning migrate into people’s intuitions, and become habitualized.”

Moreover, if this is an indication of Cappelen not denying that non-inferential mental states are treated as evidence, but suggesting an alternative account of such mental states, then it is not clear how deep the disagreement goes. As I mentioned above, friends of centrality disagree on the right account of such mental states, but they agree that such mental states are treated as evidence and refer to them as “intuitions.”

---

86 See e.g. Huemer 2005; Kaspar 2011; Parfit 2013, p. 544.
87 Cappelen 2012, p. 162; see also Deutsch 2017, p. 425.
89 Cappelen 2014a, p. 580.
90 Ibid.
91 See also Nado 2016, p. 796, n. 720.
92 See e.g. Greene 2017; Railton 2014; Sauer 2017; Woodward and Allman 2007.
Finally, my argument is not overly reliant on intuition-terminology. Intuition-terminology is combined with abduction. When both intuition-terminology and the abduction features are present, I do not find it plausible that the intuition-terminology can easily be reinterpreted in a way that suggests it is not used to denote the relevant class of mental states. For example, Cappelen suggests that intuition-terminology can be used as hedge terms—that is, terms that a speaker uses to weaken her commitment to a claim. However, given that the claim in question is treated as a starting point for an abductive argument, it seems strange to assume that Dworkin and Cohen are weakening their commitment to it. The presence of the abduction features suggests that they are committed more strongly to the case verdicts than to the theoretical accounts.

On the other hand, I think that the discussion in this section makes it clear that intuition-terminology should be understood as designating the mental state of how one is appeared-to. It provides a plausible explanation of why an author would assume that the case verdict \( p \) is shared by her peers when the case is novel and not used to illustrate a commonly accepted argument: \( p \) seems true to her; therefore, she assumes that \( p \) seems true to her peers as well.

### 8.7 Conclusion

In this article, I have investigated a hypothesis derived from centrality: intuitions served as central evidence when Dworkin and Cohen first articulated and debated luck egalitarianism. To do this, I developed an analytical framework that overcomes the positive-evidence challenge. I have argued that if case verdicts are treated as starting points for abductive arguments, it suggests that they are treated as evidence. Therefore, the framework offers diagnostics for identifying whether case verdicts are so treated. Moreover, I have argued that if case verdicts treated as starting points for abductive arguments are referred to using intuition-terminology, that gives us reason to believe that the case verdicts are treated as evidence because they are intuitive. When examining Dworkin’s and Cohen’s discussion of expensive tastes, case verdicts are treated as starting points for abductive arguments and referred to using intuition-terminology.

One particularly important lesson from this study is that we need to examine how a philosopher treats the case verdicts she reaches and how these relate to the various theoretical accounts under consideration in order to properly assess whether that philosopher is arguing for

---

a case verdict or arguing abductively. I think that insufficient attention has been paid to this point in previous studies.
The Trolley Problem and Intuitional Evidence

9.1 Introduction

It is a common assertion that the practice of treating intuitions as evidence figures centrally in philosophy. Following Herman Cappelen, we may refer to this assertion as Centrality. Recently, however, Centrality has been challenged. Most notably, Cappelen and Max Deutsch have each written a book arguing that Centrality is false. They investigate Centrality by carefully analyzing philosophical articles and the use of cases in defending and challenging theories. Roughly, they find that philosophers argue for claims about cases, which they take to suggest that Centrality is false. The idea is that since claims about cases are supported by arguments, it is not clear that the intuitiveness of the claim is doing any work. Unless friends of Centrality can demonstrate that the intuitiveness of the claims is (also) doing some epistemic work, we should assume that the claims are supported by arguments (full stop). Thus, the onus is on friends of Centrality to show that intuitions actually do epistemic work.

In this paper, I take on this challenge by combining three strategies for examining Centrality. I apply these strategies to Judith J. Thomson’s article on the trolley problem, which is an article also examined by Cappelen. The first strategy is to examine how philosophers...
argue. According to the framework I apply, if we want to know whether a certain claim is treated as evidence because of its intuitiveness, we should first consider whether it is treated as a starting point for an abductive argument. Second, if it is so treated, we should move on to consider whether intuition-terminology is used to refer to the claim. If that is the case, we have reason to believe that the claim is treated as evidence because of its intuitiveness. The second strategy is to look at what an author states to be her evidence for a particular theoretical account. The third strategy is to look at what the philosophers write about their methodology.

Based on the findings of the analysis, I argue that, contrary to what Cappelen believes, Thomson treats intuitions about cases as evidence. Moreover, I argue that the analysis reveals that previous attempts at analyzing how philosophers argue (such as Cappelen’s analysis) have considerable shortcomings. When applying the strategies of those previous studies, one runs the risk of not capturing the central features of intuitionistic methods—or even mistakenly interpreting these features as suggesting that Centrality is false.

I proceed as follows: First, I outline Cappelen’s general project. Second, I present Cappelen’s analysis of Thomson’s article and some objections to his analysis. Third, I outline how Thomson describes her method. The section on Thomson’s method serves various purposes. It brings to the fore the striking contrasts between how Thomson describes her method and what Cappelen argues is her method and thus motivates a reanalysis of it. Moreover, Thomson’s description of her method provides the basis of the third strategy I apply in my analysis of her argumentation in her article on the trolley problem. Fourth, I analyze Thomson’s article, applying the three strategies described above. Fifth, I argue that my analysis reveals some shortcomings of Cappelen’s analysis of Thomson’s article. Finally, I discuss some objections to my analysis and conclude.

9.2 Backdrop

Cappelen’s analysis of Thomson’s article is part of a bigger metaphilosophical project. Cappelen’s aim is to correct an assertion—i.e., Centrality—that “is almost universally accepted in current metaphilosophical debates and figures prominently in our self-conception as analytic philosophers.” Centrality has made philosophers engaged in first-order philosophy concerned about their disciplines as “they are unsure what intuitions are and whether they can

---


6 Cappelen, Philosophy without Intuitions, op. cit., 1.
carry the evidential weight we allegedly assign to them.”⁷ Cappelen argues that Centrality is false and hence, that this concern is unwarranted.

Cappelen’s project is wide-ranging, and his goal is to show that Centrality is false for analytic philosophy in general. He examines cases from a variety of philosophical disciplines, such as philosophy of mind and language, epistemology, and moral philosophy. The cases from Thomson’s article on the trolley problem are among those from moral philosophy⁸ and are, according to Cappelen,⁹ the best illustration of how Centrality fundamentally misrepresents the philosophical practice. The aim of Thomson’s article is to provide a solution to what she calls the trolley problem: why may a bystander divert a trolley onto a track where it will kill one person to save five, while a surgeon may not take the organs from a young man, and thereby killing him, to save five people?

9.3 Cappelen and the Trolley Problem

Cappelen starts his discussion of Thomson’s article by presenting two of the cases she considers: Trolley driver and Transplant. In Trolley driver, the only way for a trolley driver to avoid running over five people, and thereby killing them, is to divert the trolley onto another track where he will run over and kill one person. In Transplant, the only way for a surgeon to save five patients in need of new organs is to take the organs from a sixth patient and thereby kill him. These cases were first presented in Foot’s article “The Problem of Abortion and the Doctrine of Double Effect.”¹⁰ Cappelen then notes how Thomson responds to these cases. She claims that everyone to whom she has presented the cases has responded that it is permissible for the driver to divert the trolley, but impermissible for the surgeon to operate. Cappelen then goes on to claim that Thomson, after having presented the cases, “immediately goes on to question these answers. She asks for reasons (i.e. arguments) for why these should be the correct answers.”¹¹ To support this claim, he provides the following quote:¹²

Here then is Mrs. Foot’s problem: “Why is it that the trolley driver may turn his trolley, though the surgeon may not remove the young man’s lungs, kidneys, and heart?” In both cases, one will die if the agent acts, but five will live who would otherwise die—a net

---

⁷ Ibid.
⁸ He also examines Thomson’s (“A Defense of Abortion,” Philosophy and Public Affairs 1, no. 1 (1971): 47-66) violinist case. However, I do not have space to examine that case in this paper.
⁹ Cappelen, Philosophy without Intuitions, op. cit., 160.
¹¹ Cappelen, Philosophy without Intuitions, op. cit., 159-160, italics in original.
¹² Ibid., 160.
saving of four lives. What difference in the other facts of these cases explains the moral difference between them?\textsuperscript{13}

Thus, according to Cappelen:\textsuperscript{14}

A central goal of Thomson’s paper is to question the answers given to questions asked about the cases. There is a tension between the two answers and that tension shows further reflection and investigation is needed. Thomson and Foot articulate the reasons that can be given for the initial answer.

He then presents what he understands to be Thomson’s first take at how the claims about the cases can be justified by quoting a selection of passages from the article:

\begin{quote}
Killing one is worse than letting five die.
So the surgeon must refrain from operating.
By contrast, the trolley driver’s choice is between turning the trolley, in which case he kills one, and not turning the trolley, in which case he does not let five die, he positively kills them.
Now surely we can say
Killing five is worse than killing one.
But then that is why the trolley driver may turn his trolley: He would be doing what is worse if he fails to turn it, since if he fails to turn it he kills five.\textsuperscript{15}
\end{quote}

The latter quote is what Cappelen takes to be Thomson’s arguments supporting the claims about Trolley driver and Transplant.

Cappelen then goes on to apply his analytical framework. He uses three diagnostics for identifying reliance on intuition as evidence. First, intuitive judgments have a special phenomenology; they seem true (F1). Second, intuitive judgments have a special epistemic status—“Rock,” as Cappelen calls it (F2). This means that intuitive judgments have non-inferential justification and that they are evidence recalcitrant. Third, intuitive judgments are based solely on conceptual competence (F3). The absence of all three features in an argument, Cappelen thinks, “is very strong evidence that there is no reliance on the intuitive in [an] argument.”\textsuperscript{16}

\textsuperscript{13} Thomson, “The Trolley Problem,” op. cit., 1396.
\textsuperscript{14} Cappelen, Philosophy without Intuitions, op. cit. 160, italics in original.
\textsuperscript{15} Ibid. The passages are collected from Thomson, “The Trolley Problem,” op. cit., 1396-97.
\textsuperscript{16} Ibid., 114.
Cappelen thinks that the presence of F1 and F3 are indicated by an author explicitly stating that a claim has the relevant feature—that is, it either has a special phenomenology or is delivered by conceptual competence. When it comes to F2, however, he looks for evidence that is more implicit. Roughly, he thinks that a philosopher arguing for a claim gives us reason to believe the claim lacks the “Rock” status.

Cappelen’s analysis of F1 and F3 is relatively brief. He more or less just remarks that Thomson makes no statements suggesting that the features are present. His discussion of F2, on the other hand, is lengthier. In short, he maintains that Thomson argues for the claim about the cases (see the quoted passage above). I will, following Jennifer Nado, refer to the view that Thomson argues for the claims about the cases and that this suggests that intuitions about the cases are not treated as evidence as the “argument from argumentation.”

The argument from argumentation has been challenged by friends of Centrality. It has been objected that the presence of arguments need not be Centrality unfriendly. For example, arguments may be used as further support, or they may be abductive. If the arguments are abductive, the aim is to explain the intuition, and insofar as a theoretical account can explain the intuition, the theoretical account is considered plausible. Thus, the direction of support is the opposite of what Cappelen suggests; it goes from the claim about the case to a theoretical account.

Cappelen has, in response to these objections, claimed that it is not enough to argue that the presence of arguments need not be Centrality unfriendly. To push the debate forward, friends of Centrality need to come up with a method capable of providing positive evidence for Centrality—i.e., they have to show that the intuitiveness of a claim is actually doing epistemic work.

Moreover, there is a further complication that friends of Centrality have to deal with. According to Cappelen, finding claims that are not argued for is not positive evidence for Centrality. In fact, Cappelen finds some claims that Thomson does not argue for. He asserts

---

that these claims are common ground and that there is therefore no need to appeal to their intuitiveness.\textsuperscript{21} Furthermore, he argues that we should resist the idea that the claim happens to be in the common ground because of its intuitiveness: notions such as “Rock” are elusive and controversial categories, whereas “common ground” is not.\textsuperscript{22}

I think it is debatable whether Cappelen succeeds in establishing that the burden of evidence is on friends of Centrality.\textsuperscript{23} Nevertheless, Cappelen’s challenges suggest some issues that warrant consideration. This paper offers some responses to these challenges.

\subsection*{9.4 Thomson on Her Method}

In her writings, Thomson is quite explicit about her reliance on an intuition-based method.\textsuperscript{24} She writes, for example, that a moral theory “is at the mercy” of moral intuitions about particular examples.\textsuperscript{25} Elsewhere, she writes that such judgments “serve as data for” moral theories.\textsuperscript{26} In yet another piece, she writes that

\begin{quote}
[...]
the moral theorist must attend to his or her own moral beliefs about examples, stories, and cases, actual or invented, looking to see how those beliefs do (or do not) change as the details of the stories are altered. For it is precisely those beliefs which supply the data for moral theorizing, and which go a long way—if not all the way—to setting the constraints on what constitutes an acceptable moral principle, and thus on what constitutes an acceptable way of understanding what we ourselves take morality to require of us.\textsuperscript{27}
\end{quote}

From these quotes, I think it is quite clear that Thomson claims to be treating intuitions as evidence. She explicitly refers to them as “data.” She claims that moral theories are at the mercy of intuitions and that intuitions set the constraints on what constitutes an acceptable theoretical account. Thus, theoretical accounts, in Thomson’s view, should conform to our moral intuitions.

Moreover, she claims that the method she applies

\textsuperscript{22} Ibid., 154-155.
\textsuperscript{23} See e.g. Nado, “The intuition deniers,” op. cit.
\textsuperscript{24} Thomson uses “belief,” “judgment,” and “intuition” interchangeably when referring to claims about cases.
[...] is like the process Rawls describes as the effort to reach ‘reflective equilibrium’, with this proviso: on Rawls’ account of the matter, everything is provisional, everything is open to revision, whereas I am suggesting that some moral judgments are plausibly viewed as necessary truths and hence not open to revision.²⁸

In this passage, she not only claims that intuitions are evidence but that they are very strong evidence. Some intuitions even are not open to revision.

To sum up, Thomson explicitly states that she treats intuitions about cases as “data.” Moreover, some intuitions are treated as very strong evidence, as they are not open to revision. Interestingly, Thomson’s description of her method starkly contrasts Cappelen’s findings. Thus, it is worth taking a new look at Thomson’s article.

9.5 A Reanalysis of Thomson’s Article

In this section, I do a reanalysis of Thomson’s article and apply several strategies to this endeavor. First, I apply an analytic framework developed in Anonymous.²⁹ Second, I look at what Thomson states is her evidence for or against the theoretical accounts she discusses in her article. Third, I consider what Thomson writes about the method she claims to apply.³⁰

²⁸ Thomson, The Realm of Rights, op. cit. (1990), 32 n. 20.
²⁹ Conte, op. cit.
³⁰ Other counter-readings of Cappelen’s cases have been offered. Chalmers, for example, provides a counter-reading of his own “zombie argument.” Chalmers, op. cit. Chalmers claims that the observations that Cappelen takes to undercut the support for Centrality are in fact consistent with Centrality. Ibid., 540-541. While Chalmers mentions some possible sources of positive evidence for Centrality, Cappelen argues that Chalmers falls short of providing any positive evidence for Centrality. Cappelen, Philosophy without Intuitions, op. cit. (2012), 584-587. For example, Chalmers mentions that there is extensive intuition-talk in philosophical texts and that many philosophers claim that they rely on intuitions as evidence. However, he does not engage with Cappelen’s objections to the strategy of treating intuition-talk as evidence for Centrality. Nor does he engage with Cappelen’s claim that philosophers have false second-order beliefs. Thus, while Chalmers argues that Cappelen’s observations are consistent with Centrality, my strategy is to go beyond Cappelen’s observations and investigate if further observations about what Thomson is doing allow us to discriminate between Cappelen’s (Centrality unfriendly) reading of the text and a Centrality supporting reading of the text. Moreover, I engage with Cappelen’s arguments concerning intuition-talk and second-order beliefs. Thanks to a reviewer for the journal for encouraging me to distinguish my strategy from Chalmers’.

In addition, it is worth mentioning that Kevin Tobia offers a counter-reading of Thomson’s article. Tobia, “Philosophical Method and Intuitions as Assumption,” Metaphilosophy 46, nos. 4-5 (2015): 575-594. However, Tobia’s analysis of Thomson’s argument is rather short, and he does not explicitly engage with the challenges noted in Section 3 of this paper. That said, Tobia’s reading is similar to the one I offer.
9.5.1 Arguments

The analytic framework consists of two steps. The first step is to find out whether claims about cases are treated as starting points for abductive arguments. This step is in line with some of the responses to Cappelen, suggesting that many philosophical arguments are in fact abductive, which means that the direction of support goes from claims about cases to theoretical accounts. However, the framework expands on this strategy by offering a set of diagnostics that help determine whether the direction of support goes from the claims to the theoretical accounts or the other way. The diagnostics, which I will refer to as the “abduction features,” are as follows:

When a claim about a case, c, is treated as a starting point for an abductive argument:

(i) Not-c is typically not a salient option.
(ii) Several theoretical accounts are under consideration.
(iii) A theoretical account is considered more/less plausible if it can/cannot explain c, all else being equal.

(i) indicates that the point of the inquiry is not to find out how to judge the case and that c is taken for granted. If, on the contrary, not-c is a salient option, this could suggest that the aim of the discussion of the case is to make (and justify) a judgment about the case. This provides us with reason to believe that the argument in question is not an abductive one. Moreover, if only one theoretical account is under consideration, this provides us with further reason to believe that the aim of the discussion is to judge the case, as the theoretical account seems to be taken for granted. Thus, this suggests that the theoretical account is used to eliminate or support the case verdict. This is not to say that the purpose of philosophical debates is to provide answers to questions about specific cases. Rather, the cases are used to raise interesting questions, the answers to which inform the more general philosophical question under debate. The presence of (ii), however, indicates that one is trying to decide which theoretical account is more plausible. To be more specific, the idea is that when there are several theoretical accounts that

\[31\] Conte, op. cit.
\[32\] The original framework also includes a fourth feature (roughly, that claims about new cases are introduced as arbiters when several competing theoretical accounts can explain c), but I have not included it here since it is not necessary to show that Thomson treats case verdicts as starting points for abductive arguments. It suffices to show that (i), (ii), and (iii) are present.
\[33\] Cappelen, Philosophy without Intuitions, op. cit., 122-123.
\[34\] Ibid., 188.
apply to the case, it gives us reason to believe that the theoretical accounts are under consideration and are not provided as support for c.\footnote{Thanks to a reviewer of the journal for pressing me to clarify these points.}

However, the presence of (i) and (ii) need not suggest that c provides the theoretical account with any (theory) independent credibility; for example, the support could go in both directions.\footnote{Cappelen, \textit{Philosophy without Intuitions}, op. cit., 123; Max Deutsch, “Replies to commentators,” \textit{Inquiry} 60, no. 4 (2017): 420-442, 425.} However, the presence of (iii) indicates that c has some credibility independent of the theoretical account in virtue of which it confers credibility onto the theoretical account. The presence of (i), (ii), \textit{and} (iii), then, suggests that the claim is used as a starting point for an abductive argument and thus is treated as evidence.

However, cf. Cappelen, it is possible that some claims are so treated because they are common ground for some extra-intuitive reason. Thus, we need a strategy for finding out whether the claim is used as a starting point \textit{because} of its intuitiveness. This brings us to the second step of the method: If intuition-terminology is used to describe the claim, then we have reason to believe that the claim is used as a starting point because of its intuitiveness.\footnote{Since the framework does not treat the presence of intuition-terminology by itself as a sufficient condition for believing that intuitions are treated as evidence, the framework avoids the objections against relying on the use of intuition-terminology. See e.g. Cappelen, \textit{Philosophy without Intuitions}, op. cit., Part 1. For a more complete discussion of that issue, see Conte, op. cit. See also the discussion in Section 7.2.} Here are some examples of such intuition-terminology:

\begin{quote}
“Intuition,” “intuitive(ly),” “seem,” “look,” “see,” “what we would say,” “feel,” “clear(ly),” and “obvious(ly).”\footnote{Bengson, op. cit., 561.}
\end{quote}

To sum up, the method can be portrayed as a decision tree (see figure 1). First, we ask whether the abduction features are present. If “no,” then the claim is not a starting point of an abductive argument. If “yes,” we go on to consider whether intuition-terminology is used when making the claim or when referring to it. If we do not find such terminology, this could suggest that the claim is common ground for some extra-intuitive reason (or that the source of the claim is not mentioned). If we find such terminology, we have reason to believe that the claim is treated as evidence because of its intuitiveness.\footnote{This approach may seem to be similar to that of James Andow. James Andow, “Abduction by Philosophers: Reorienting Philosophical Methodology,” 47, no. 3 (2016): 353-370. However, Andow does not focus on abductive arguments as such. Ibid., 362. He suggests that Cappelen only denies “that philosophers \textit{rely} on intuitions \textit{as construed in certain very particular ways} as evidence in a \textit{distinctive way}.” Ibid., 357, italics in}

---

\footnotesize

35 Thanks to a reviewer of the journal for pressing me to clarify these points.
37 Since the framework does not treat the presence of intuition-terminology by itself as a sufficient condition for believing that intuitions are treated as evidence, the framework avoids the objections against relying on the use of intuition-terminology. See e.g. Cappelen, \textit{Philosophy without Intuitions}, op. cit., Part 1. For a more complete discussion of that issue, see Conte, op. cit. See also the discussion in Section 7.2.
38 Bengson, op. cit., 561.
39 This approach may seem to be similar to that of James Andow. James Andow, “Abduction by Philosophers: Reorienting Philosophical Methodology,” 47, no. 3 (2016): 353-370. However, Andow does not focus on abductive arguments as such. Ibid., 362. He suggests that Cappelen only denies “that philosophers \textit{rely} on intuitions \textit{as construed in certain very particular ways} as evidence in a \textit{distinctive way}.” Ibid., 357, italics in
Fig. 1 The analytic framework (Figure from Conte, op. cit.)

9.5.1.1 The abduction features

Not-c is typically not a salient option: When Thomson presents cases describing actions, it is almost always the case that the actions “seem,” “feel,” “obviously are,” or “clearly are” either permissible or impermissible (c). The opposite claim (not-c) is not considered an option. For example, the possibility that it is permissible for a surgeon to take the organs from a healthy patient to save five others is never under consideration.

Some of the theoretical accounts that Thomson considers have not-c as an implication. Nevertheless, this does not make Thomson consider not-c a salient option. If a theoretical account has implications that conflict with a claim about a case, her reaction is to revise the theoretical account. Thus, if a theoretical account has not-c as an implication, this is considered as something that counts against the theoretical account and may call for revision or even rejection of it, but not as a reason to consider not-c.

Here is an example: Thomson presents the following theoretical account from Foot: “(II) Killing five is worse than killing one.”40 To assess this theoretical account, she considers alternate versions of the Transplant. In these versions, the surgeon is responsible for the five persons needing new organs to survive, either due to carelessness or because he wanted to kill them (but later repents). To the question of whether the surgeon may cut open the one and

---


original; but see Cappelen, Philosophy without Intuitions, op. cit., 3, 18. I focus specifically on the question of whether intuitions are treated as starting points for abductive inferences and argue that Cappelen’s interpretation of Thomson’s arguments is wrong. Thus, my disagreement with Cappelen cannot be reduced to a disagreement about how to understand Centrality.
distribute his organs to the five, Thomson responds, “I should think plainly not.”

She then goes on to consider how that claim reflects on (II):

There are two ways in which we can go here. First, we can say: (II) does tell us that the surgeon ought to operate, and that shows it is false. Second, we can say: (II) does not tell us that the surgeon ought to operate, and it is true.

What she is claiming, then, is that the truth of (II) is contingent on whether it is consistent with the claim (c) about the alternate version of the Transplant. If the implication of (II) is not-c, then it is “false.” If the implication of (II) is c, then it is “true.”

Several theoretical accounts are under consideration: Thomson considers several theoretical accounts. For example, in addition to (I) and (II), she considers a revised version of (II):

(ii’) If a person is faced with a choice between doing something here and now to five, by the doing of which he will kill them, and doing something else here and now to one, by the doing of which he will kill only the one, then (other things being equal) he ought to choose the second alternative rather than the first.

Moreover, she considers the Kantian idea that you should always treat persons as an end and never as a means and the idea that “it is not morally required of us that we let a burden descend out of the blue onto five when we can make it instead descend onto one if we can make it descend onto the one by means which do not themselves constitute infringements of rights of the one”.

A theoretical account is considered more/less plausible if it can/cannot explain c: A clear example of a theoretical account being considered plausible if it can explain claims about cases, and implausible if it cannot, is the example mentioned above when discussing feature (i). (II) is true if it can explain c, but false if it cannot.

Another example is Thomson’s discussion of the first part of Foot’s principle—“(I) Killing one is worse than letting five die.” In that discussion, she introduces the bystander case

---

41 Ibid., 1399.
42 Ibid., 1400.
43 Ibid., italics in original.
44 Ibid., 1401.
45 Ibid., 1409, italics in original.
46 Ibid., 1396.
(henceforth, Bystander): a bystander may throw a switch, and thereby turn a trolley that otherwise would have killed five persons onto a track where it will only kill one. Thomson’s “feeling” is that the bystander may throw the switch. She continues by saying that “If [the bystander] may [throw the switch], there is serious trouble for Mrs. Foot’s thesis (I)” and that “If thesis (I) were true, it would follow that the bystander may not throw the switch, and that I am taking to be false.” Hence, since (I) is in conflict with the feeling about the switch case, which she takes to be true, (I) is considered to be less plausible than it otherwise would have been. Consequently, Thomson goes on to search for other theoretical accounts that are consistent with the claims about Bystander and the other cases.

To sum up, Thomson’s approach is the following: She starts with a set of cases and a theoretical account that is supposed to account for these cases and that she wants to investigate. To investigate the accounts, she constructs new cases. In these cases, there is only one salient option (i). If the account is inconsistent with a claim about a case, she revises the theoretical account and further investigates it by constructing new cases ([ii] and [iii]). She does this until she comes up with an account she deems satisfactory. The presence of features (i), (ii), and (iii), then, suggests that Thomson is treating claims about cases as starting points for abductive arguments.

### 9.5.1.2 Intuition-terminology

Thomson consistently refers to claims about cases using intuition-terminology. Sometimes she refers to how she or people in general “feel” about a case and how things “seem.” Other examples of intuition-terminology are “plain(ly),” and “obvious” and “clear.” Moreover, when discussing what she calls the “distributive exemption,” she explicitly refers to such claims as intuitions:

I do not find it clear why there should be an exemption for, and only for, making a burden which is descending onto five descend, instead, onto one. That there is seems to me very plausible, however. On the one hand, the agent who acts under this exemption makes be a threat to one something that is already a threat to more, and thus something that will do harm whatever he does; on the other hand, the exemption seems to allow

---

47 Ibid., 1397, 1398.
48 Ibid., 1398-1399.
49 Ibid., 1397, 1398, 1403, 1412.
50 Ibid., 1398, 1403, 1408.
51 Ibid., 1399, 1407.
52 Ibid., 1411.
those acts which *intuition* tells us are clearly permissible, and to rule out those acts which *intuition* tells us are clearly impermissible.\(^53\)

What Thomson is saying is that the fact that the exemption is consistent with our intuitions about acts counts in favor of the exemption’s plausibility. Note that the acts which “intuition” tells us are permissible or impermissible are the acts depicted in the cases Thomson discusses—e.g. Transplant, Trolley driver, and Bystander.\(^54\)

In sum, that Thomson treats claims about cases as starting points for abductive arguments and refers to them using intuition-terminology suggests that she treats the claims as evidence because of their intuitiveness.

### 9.5.2 Explicit Statements

A second strategy is to look at what an author states is her evidence for or against the theoretical accounts she discusses in her article. For example, we might look for whether an author claims she is engaging in abductive reasoning and at what she claims is the starting point of such reasoning. Cappelen\(^55\) thinks that looking for such statements is an “obvious” way of approaching the question of whether an author engages in abductive reasoning. Interestingly, Cappelen\(^56\) claims that “in none of the cases [he investigates] is there any explicit statement like that.” In this section, I examine several passages from Thomson in which I argue there are, contrary to what Cappelen claims, such statements. Moreover, I also look for explicit statements of intuitions (about cases) being treated as evidence for, or reasons for accepting, a theoretical account.

#### 9.5.2.1 Abductive Arguments

Let us start with considering whether Thomson is explicitly stating that she engages in abductive reasoning. When arguing abductively, one treats some claims as datum and infers

---

\(^{53}\) Ibid., 1408, Thomson’s italics omitted, mine added.

\(^{54}\) A critic might want to point out that when Thomson is commenting on the Transplant and Trolley she writes that “[e]verybody to whom I have put this […] case says” Yes/No. Ibid., 1395–1396. These statements could suggest that Thomson has a fairly good idea about what her peers think about the cases and is possibly appealing to these claims because they are in the common ground (and that they are in the common ground for all kinds of different reasons, but not because they are intuitive). First, formulations such as “what people would say” about a case can be used to indicate what intuitions people have about a case. Second, I do find it hard to interpret the passage cited in this section in a way that supports this reading of Thomson, in particular when taking into account the second and third strategies of my analysis.

\(^{55}\) Cappelen, *Philosophy without Intuitions*, op. cit., 122.

\(^{56}\) Ibid.
that the hypothesis (derived from, for example, a theoretical account) that best explains the datum is true.  

Thomson explicitly calls for theoretical accounts that can explain the claims about the cases under consideration. Here are a couple of examples:

1. … I propose we grant that both the bystander and the surgeon would infringe a right of their ones, a right in the cluster of rights that the ones’ have in having a right to life, and that we look for some other difference between the cases which could be appealed to to explain the moral difference between them.

2. As I shall put it, shoving a person, toppling a person off a footbridge [i.e., what the person in the footbridge case must do to save the five], are themselves infringements of rights of his. A theory of rights ought to give an account of what makes it be the case that doing either of these things is itself an infringement of a right of his. But I think we may take it to be a datum that it is, the job which confronts the theorist of rights being, not to establish that it is, but rather to explain why it is.

The first quote is preceded by a discussion of whether the morally relevant difference between Bystander and Transplant is that the bystander would not infringe a right of the one on the other track by diverting the trolley, whereas the surgeon would infringe a right of the sixth patient by operating. However, she argues that both the bystander and the surgeon infringe a right of the persons in question by proceeding, and thus that this difference cannot explain the claims about the cases—i.e., that the bystander may proceed, whereas the surgeon may not. I think, therefore, that by using “explain” in the second quote, Thomson is signaling that she is engaging in abductive reasoning and treating the claims about the cases as datum.

In the second quote, Thomson explicitly states that she treats the claim that “shoving a person, toppling a person off a footbridge are themselves infringements of a right of [the person]” as “datum.” Moreover, she claims that the job that confronts us is not to establish such claims, but to provide a theoretical account that can explain them. The second quote, then, leaves little room for doubting that Thomson is stating that she is engaging in abductive reasoning. (Note that the point of explaining the datum is not to shore up the claim treated as datum. Thomson explicitly states that our job is not to establish this (or these) claim(s).)

58 See e.g. Thomson, “The Trolley Problem,” op. cit., 1396, 1400, 1401, 1404, 1405, 1409, 1410.
59 Ibid., 1406, italics in original.
60 Ibid., 1409, Thomson’s italics omitted, mine added
9.5.2.2 The role of intuitions

When considering whether Thomson states that intuitions serve as evidence for or against a claim, the paragraph in which Thomson considers the distributive exemption is worth having a closer look at:

I do not find it clear why there should be an exemption for, and only for, making a burden which is descending onto five descend, instead, onto one. That there is seems to me very plausible, however. On the one hand, the agent who acts under this exemption makes be a threat to one something that is already a threat to more, and thus something that will do harm whatever he does; on the other hand, the exemption seems to allow those acts which intuition tells us are clearly permissible, and to rule out those acts which intuition tells us are clearly impermissible.61

In this paragraph, Thomson mentions two reasons that make it plausible that there is an exemption for making a burden that is descending onto five descend, instead, onto one. It is the second reason that is of particular interest. Thomson is claiming that the exemption allows those acts which intuition tells us are clearly permissible and rules out those acts which intuition tells us are clearly impermissible is a reason to think that the exemption is plausible. This, then, is an example of Thomson explicitly stating that she treats intuitions as evidence.

9.5.3 Beliefs about Methodology

A third strategy is to consider what an author has said or written about her methodology. According to this strategy, when we want to know what method an author applies in a particular text, we can consult other texts by the author to make inferences about the method in the text we are interested in. In Section 9.4, I outlined Thomson’s methodological views, drawing on several sources. The outline in Section 9.4 revealed that Thomson believes that we should treat intuitions as evidence, and this therefore gives us reason to believe that she treats intuitions as evidence.

9.6 Explaining Why Cappelen is Wrong

Above, I pursued several strategies in trying to answer the question of whether Thomson treats intuitions as evidence. The strategies unanimously answer the question in the affirmative. Although I think my analysis provides a strong defense of Centrality on its own, it would further

61 Ibid., 1408, my italics.
strengthen it if I were able to account for the findings of Cappelen’s study and explain where his analysis goes wrong. That is what I attempt to do in this section.

First, Cappelen claims that Thomson “asks for reasons.” In Section 9.5.2.1, I hinted at a response to this claim. It is worth noting that Thomson does not ask for reasons but rather explanations. Thomson’s request, then, is by the very least, not something that counts against the assumption that Thomson engages in abductive reasoning. Thus, a shortcoming of Cappelen’s study is that he fails to take seriously the possibility that what he assumes are reasons supporting claims are rather explanations of them.

Second, as I outlined in Section 9.3, Cappelen provides a suggestion for what we might think is an argument that Thomson provides for the claims about Transplant and Trolley driver. However, my analysis suggests that Thomson does not deduce the claims from the theoretical accounts, but rather treats the claims as starting points for abductive arguments in which the acceptance of the theoretical accounts figures as the conclusion. Nothing in the passage that Cappelen cites contradicts my interpretation of Thomson’s arguments. However, some of the findings of my study contradict the view that Thomson is arguing for the claims. In particular, the presence of the abduction features strongly suggests that the theoretical accounts she is discussing are not used to support the claims about the cases. Rather, the direction of support is the opposite. For example, the first premise in what Cappelen takes to be Thomson’s argument, as argued in Section 9.5.1.2, is a theoretical account that Thomson rejects because it cannot explain the claim about Bystander. Moreover, although what apparently supports the claim about Transplant is rejected, Thomson does not weaken her commitment to the claim. Thus, a shortcoming of Cappelen’s study is that his analysis is too narrow in the sense that he considers too few cases and theoretical accounts to get a proper grasp of the dialectic of Thomson’s article.

9.7 Objections and Replies

In this section, I consider some objections to my analysis. First, I consider objections to the second and third strategy. Then, I consider objections to the first strategy.

9.7.1 Philosophers Have False Second-Order Beliefs

Several intuition-deniers have claimed that we should not put too much emphasis on philosophers’ self-analyses.\(^\text{62}\) If they are right, that might seem to pose a problem for the current

analysis. Both the second and third strategy heavily rely on philosophers’ self-analyses. Let us, therefore, consider the arguments used to support this claim in more detail.

Both Cappelen\(^{63}\) and Deutsch\(^{64}\) mainly support the claim that philosophers have false second-order beliefs about their methods by appealing to the findings of their case studies. However, Deutsch does not analyze Thomson’s arguments, and I have argued that the findings of Cappelen’s analysis are contestable.

Molyneux, on the other hand, offers some arguments that do not depend on the findings of such analyses. Molyneux\(^{65}\) claims that “individuals are often unaware of what is truly guiding their practice, and tend to manufacture post hoc rational motivations when challenged,” and he cites a study by Richard E. Nisbett and Timothy DeCamp Wilson\(^{66}\) to support the claim. Moreover, he thinks that the fact that “epistemologists prior to Gettier (1963) mistakenly thought that they were disposed to apply the term ‘knowledge’ only to cases of justified true belief” illustrates that “philosophers should already be aware of the dangers of privileging their own self-analyses.”\(^{67}\) Finally, he maintains that “[u]nless there is some good reason why we might misanalyze our own application of a central term yet not be prone to misanalyzing our own application of a central method, we must be cautious about placing too much weight on our own supposed self-knowledge.”\(^{68}\)

It is not clear to me to what extent Molyneux’s observation about epistemologists’ application of the term “knowledge” compels us to doubt philosophers’ self-analyses. Linda Zagzebski\(^{69}\) claims that “[t]he moral drawn in the thirty years since Gettier published his famous paper is that either justified true belief (JTB) is not sufficient for knowledge, in which case knowledge must have an ‘extra’ component in addition to JTB, or else justification must be reconceived to make it sufficient for knowledge.” This does not entail that the epistemologists’ self-analyses were completely off the mark.

Moreover, I doubt that Nisbett and Wilson’s study provides us with a reason to believe that we should not trust philosophers’ self-analyses of their methods. To explain why I do so, it

---

\(^{63}\) Cappelen, *Philosophy without Intuitions*, op. cit., 163.


\(^{65}\) Molyneux, op. cit., 163.


\(^{67}\) Molyneux, op. cit., 445

\(^{68}\) Ibid.

would be helpful to consider some of the studies they conduct. Nisbett and Wilson\textsuperscript{70} conducted two studies in which participants were invited to evaluate an array of articles of clothing. Unbeknownst to the participants, the articles were identical. The authors found that the position of the object influenced the participants’ evaluation of the articles of clothing. When asked which article of clothing was the best quality, the right-most object in the array was heavily over-chosen. When asked about the reasons for their choices, however, none of the participants mentioned the position of the article in the array. Moreover, when explicitly asked about whether the position of the article may have influenced their choice, “virtually all subjects denied it…”\textsuperscript{71}

When comparing epistemologists and the participants in Nisbett and Wilson’s studies, there are some significant differences. While the participants are clueless with regard to what factors influence their reasoning, epistemologists seem to be able to identify (at least) some important characteristics of the cases to which they apply the term “knowledge.” Second, while the participants in Nisbett and Wilson’s studies doubted or denied that the stimuli they were subjected to influenced their judgments or behavior, the epistemologists did not deny the shortcomings of the JTB account of knowledge when confronted with the cases presented by Gettier. Moreover, even if epistemologists were—or are, for that matter—unaware of what makes them inclined to think that “knowledge” applies to a certain case, I do not see why it should follow that they are not able to describe the method they apply to find out when or why “knowledge” applies.\textsuperscript{72}

### 9.7.2 Objections Concerning the Analytic Framework

Some critics may have doubts concerning the analytical framework I applied as a part of the first strategy for testing Centrality. For example, it has been argued that showing that claims about cases are treated as starting points for abductive arguments does nothing to refute the

\textsuperscript{70} Nisbett and Wilson, op. cit., 243–244.

\textsuperscript{71} Ibid., 244.

\textsuperscript{72} Perhaps an objector might try to challenge the reliability of philosophers’ second-order beliefs by pointing out that Thomson claims, as quoted above, that “some moral judgments […] are not open to revision,” but in one of her recent articles, she revises the judgment about the switch case. Thomson, “Turning the Trolley,” \textit{Philosophy \\& Public Affairs} 36, no. 4 (2008): 359-374. She no longer thinks that it is permissible to turn the trolley in that case. This is an interesting objection, but I do not have space to discuss it here as it raises bigger methodological questions, such as to what extent one should analyze articles in isolation or in relation to other articles by the same author, or as part of a greater debate. So, while such a question deserves thorough discussion, I can only offer some preliminary responses to the objection here. First, it is worth noting that Thomson uses the qualifier “some” in the quote. Moreover, it could plausibly be suggested that what Thomson is doing in that article is to introduce further intuitions and that she considers the intuitions about the switch case in light of (among other things) intuitions about other cases.
argument from argumentation. The idea seems to be that when such claims are treated as starting points for—or explanandum in—abductive arguments in which a theoretical account figures as the explanans, the direction of support is likely to go both ways. When that is the case, we lack a positive reason to assume that the claims about the cases are treated as having any independent credibility, and should therefore assume that they are not treated as evidence. Thus, according to this objection, the first step of the analytic framework fails to detect whether claims about cases are treated as evidence.

Moreover, some question that intuition-terminology indicates that an author finds the claim that is described using such terminology intuitive. Instead, such terminology could be used as hedge terms—i.e., to indicate that one is weakening one’s commitment to the claim. Consequently, intuition-terminology need not indicate that claims are treated as evidence—or are in the common ground—because they are intuitive.

There are some things to be said in response to these objections. However, let us, for the sake of argument, grant the objections. Let us, moreover, again for the sake of argument, assume that philosophers’ testimonies based on self-analyses (cf. strategies 2 and 3) do not provide us with a positive reason to believe that Centrality is true. Thus, the three strategies I apply demonstrate only that it is possible that Thomson treats intuitions as evidence, but does not provide positive evidence for believing so. Thus, none of the strategies succeeds in responding to Cappelen’s challenge.

Is there anything more to be said by a friend of Centrality? I think there is. Although none of the strategies may be capable of providing positive reasons for believing that Centrality is true on their own, they might be when we combine them. If a philosopher claims that she applies a method according to which intuitions serve as evidence (cf. strategy 3), and the philosopher explicitly states that she treats a theoretical account as being able to account for our intuitions as a reason for thinking the account is plausible in her discussion of a particular first-order issue (cf. strategy 2), and if, when we examine the ways in which the philosopher actually argues, the ways she argues fit very well with what she claims (cf. strategy 1), then we should assume that she does what she says she is doing. Thus, we should assume that Thomson treats intuitions as evidence.


Conte, op. cit.
9.8 Conclusion

In this paper, I have explored three strategies for investigating Centrality. I have applied these strategies to Thomson’s article about the trolley problem. Cappelen also examined this article, considering it a good example of how Centrality misrepresents the philosophical practice. The first strategy was to examine how philosophers argue, applying an analytical framework that helps us determine whether a claim is treated as a starting point for an abductive argument, and hence evidence, because of its intuitiveness. The second strategy was to look at what an author states is her evidence in a discussion of a particular issue. The third strategy was to look at what the philosopher writes about her methodology. All three strategies yield the same conclusion: Thomson treats intuitions as evidence.

The implications of my findings are, first, that Cappelen’s approach has certain shortcomings. It does not capture the central features of intuition-based methods and even mistakenly interprets some of these features as suggesting that Centrality is false. This, I think, points to a more general critique of Cappelen’s method and thus his project more generally. Second, since Thomson treats intuitions as evidence, it is, in particular, questionable whether Cappelen manages to show that Centrality is false with regard to moral philosophy.
References


REFERENCES


Cova, Florian, Brent Strickland, Aangela Abatista, Aurélien Allard, James Andow, Mario Attie, James Beebe, Renatas Berniūnas, Jordane Boudesseul, Matteo Colombo, Fiery Cushman, Rodrigo Diaz, Noah N'Djaye Nikolai van Dongen, Vilius Dranseika, Brian D. Earp, Antonio Gaitán Torres, Ivar Hannikainen, José V. Hernandez-Conde, Wenjia Hu, François Jaquet, Kareem Khalifa, Hanna Kim, Markus Kneer, Joshua Knobe, Miklos Kurthy, Anthony Lantian, Shen-yi Liao, Edouard Machery, Tania Moerenhout,


REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


154
REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


