Olle Risberg

Guiding Concepts

Essays on Normative Concepts, Knowledge, and Deliberation
Abstract

This thesis addresses a range of questions about normativity, broadly understood. Recurring themes include (i) the idea of normative ‘action-guidance’, and the connection between normativity and motivational states, (ii) the possibility of normative knowledge and its role in deliberation, and (iii) the question of whether (and if so, how) normative concepts can themselves be evaluated.

The first two papers, ‘The Entanglement Problem and Idealization in Moral Philosophy’ and ‘Weighting Surprise Parties: Some Problems for Schroeder’, critically examine various versions of the view that what we ought to do depends on some (actual or hypothetical) motivational states, such as desires. It is suggested that such views are, for different but interrelated reasons, extensionally inadequate.

The third paper, ‘From Evolutionary Theory to Moral Skepticism, via Disagreement’ (co-authored with Folke Tersman), proposes that two arguments for moral skepticism can be combined in a mutually supportive way. A central role is played by the principle that a subject S knows that p only if S adherently believes that p, where this roughly means that S could not easily have failed to believe that p unless her epistemic position were worse or p were false. It is suggested that evolutionary considerations and facts about moral disagreement together indicate that moral beliefs violate this principle.

The fourth paper, ‘Ethics and the Question of What to Do’, offers an account of the so-called ‘central deliberative question’ that is highlighted by several kinds of choice situations, including those that involve normative uncertainty and normative conflicts. It is proposed that this question is not best understood as the question of what one ought to do, not even in an ‘all things considered’ sense, but as the question of what to do. A meta-normative view that involves elements of both cognitivism and non-cognitivism is put forward as the best explanation of this fact.

The fifth paper, ‘Meta-Skepticism’, develops a novel skeptical challenge to beliefs about the external world, the central idea being that even if beliefs about the external world can constitute knowledge, there are various other knowledge-like concepts that they cannot satisfy even if they are true. This raises the question of whether some of these concepts are epistemically more important than the others, and, in particular, the further question of how the relevant notion of ‘epistemic importance’ should be understood. Several answers to this question are considered and found wanting.

Keywords: action-guidance, adherence, all things considered ought, alternative normative concepts, conceptual engineering, debunking, deliberation, disagreement, epistemic pluralism, epistemology, ethics, external world skepticism, idealization, metaethics, meta-skepticism, moral skepticism, normativity, ought, reasons, response-dependence, uncertainty

Olle Risberg, Department of Philosophy, Ethics and Social Philosophy, Box 627, Uppsala University, SE-75126 Uppsala, Sweden.

© Olle Risberg 2020

urn:nbn:se:uu:diva-398744 (http://urn.kb.se/resolve?urn=nbn:se:uu:diva-398744)
To my sister, my mother,
and my father
List of Papers

This thesis consists of a general introduction and the following papers:

3. Risberg, Olle, and Tersman, Folke (manuscript). “From Evolutionary Theory to Moral Skepticism, via Disagreement.”
4. Risberg, Olle (manuscript). “Ethics and the Question of What to Do.”
5. Risberg, Olle (manuscript). “Meta-Skepticism.”

Essay 1 is reprinted with the kind permission from the editors of *Philosophical Quarterly*. Essay 2 is reprinted with the kind permission from Cambridge University Press.
Acknowledgements

Throughout my work on this thesis, I have received help and support from a large number of people. I am very grateful to all of them.

First of all I want to thank my main supervisor, Erik Carlson. Erik has been extremely generous with his time and his willingness to help, and his comments and suggestions have always been both insightful and constructive. Everything I have written has improved significantly because of him.

I am also indebted to my assistant supervisor, Andrew Reisner. Andrew has always been available for support and guidance, and his detailed comments on my drafts have been extremely helpful. I have learnt a lot about both writing and doing philosophy from him.

While Folke Tersman has not been my supervisor, he has in many other ways resembled one. Since 2016 we have co-operated closely on several philosophical projects, including (but not limited to) one of the essays in this thesis. The experience and knowledge that I have acquired in connection to these projects have been invaluable.

I have also been fortunate enough to co-author philosophical texts of various kinds with Per Algander, Erik Carlson, Daniel Fogal, Nils Franzén, Jens Johansson, and Victor Moberger. Working with these skilled writers has been highly instructive, and I am thankful for their willingness to do so.

I owe special thanks to Krister Bykvist, who generously gave me detailed comments on a full draft of this thesis in connection to my final seminar in September 2019. Krister was also the one who suggested the title of this thesis, ‘Guiding Concepts’, to me. For comments on the whole manuscript, I am also thankful to Jens Johansson and Victor Moberger.

During my time as a PhD student, the Department of Philosophy at Uppsala University has provided me with extensive teaching opportunities, funding for travels to various conferences and workshops, and all in all a great research environment. Matti Eklund has not only repeatedly given me valuable philosophical comments and advice, but has also helped me with several practical issues. Susanne Gauffin, Anna Gustafsson, Paulina Hansen, Johan Löfström, Rysiek Sliwinski and Ulrika Valdeson have also frequently helped me with various practical matters. Drafts of most parts of this thesis have been presented at the Uppsala Higher Seminar in Practical Philosophy, and I am thankful to the participants at these seminars for their comments and discussion. Other Uppsala philosophers who have given me valuable advice, inspiration and support include Tobias Alexius, Emil Andersson, Karl Bergman,

From September 2018 to February 2019 I was a visiting scholar at Columbia University in New York. I am thankful to Justin Clarke-Doane, my mentor at Columbia, for the opportunity to make this visit, and for his helpful comments and interesting discussions. I am also grateful to the graduate students at Columbia for making my visit there very enjoyable, and for their feedback on an early version of Essay 5 in this thesis, which was presented at the Columbia Graduate Student Workshop in November 2018.

During my time overseas, I twice had the opportunity to visit Don Loeb at the University of Vermont. At one of these occasions I was also invited to present an early version of Essay 5 at the UVM Ethics Reading Group. Thanks to the participants for their helpful comments. I am especially grateful to Don and Barbara for their hospitality and the interesting philosophical discussions.

The ideas in this thesis have been presented at several other conferences and workshops as well. In August 2017, Essay 2 was presented at the European Conference of Analytic Philosophy 9 at LMU Munich. In April 2017, some of the material that would eventually lead up to Essay 3 was presented at the 3rd Varieties of Normativity Workshop at Uppsala University, and at Evolutionary Ethics: Debunking & Disagreement at Utrecht University. In June 2018, Essay 4 was presented at the conference The Future of Normativity at University of Kent. Finally, in June 2019, Essay 5 was also presented at the Swedish Congress of Philosophy in Umeå. I am grateful to the organizers and audiences at these occasions.

During my work on this thesis I have received generous financial support from Gästrike-Hälsinge nation in Uppsala (Göransson-Sandviken’s research scholarship), Helge Ax:son Johnson’s Foundation, Hultengren’s Foundation, J. Håkansson’s Grant Foundation, The Royal Swedish Academy of Letters, History and Antiquities, The Salén Foundation, Sixten Gemzéus’ Foundation, Thun’s Grant Foundation, Uplands nation in Uppsala (V. Ekman’s scholarship), and Värmland’s nation in Uppsala (Uddeholm’s travel scholarship). I am very thankful to all of them.
Outside of academia, I am especially thankful to Sartez Faraj, from whom I have learnt incredibly much about both music and other things in life. I am also indebted to Hugo Duregård, Daniel Göransson, and Axel Schad Elgstrand for their long-lasting friendship.

Finally, I have had the enormous fortune of being part of a large and loving family. For all their support throughout the years, I am deeply grateful to my grandparents, aunts, uncles, cousins, and, in particular, to my sister Klara, my mother Susanna, and my father Mattias, to whom this thesis is dedicated.
Contents

General Introduction ........................................................................................................11
1. Introduction .............................................................................................................11
2. The Study of Normativity: Some Distinctions ......................................................14
3. Normative Facts and Desires .............................................................................16
   3.1 The Entanglement Problem ...........................................................................21
   3.2 Hypotheticalism ............................................................................................24
4. Skepticism About What We Ought to Do ...........................................................29
5. The Question of What to Do ...............................................................................34
6. Normative and Epistemic Pluralism ....................................................................38
7. Essay Summaries .................................................................................................42
References ..................................................................................................................44
1. Introduction

We often wonder what we ought to do, how we ought to feel, and what we ought to believe. These questions are examples of normative questions. While normative questions include questions about ‘oughts’, they are not limited to them; they also include questions about what is good and bad, right and wrong, moral and immoral, prudent and imprudent, virtuous and vicious, reasonable and unreasonable, and so on.\(^1\) This thesis concerns a range of issues in the study of normativity, broadly understood. It consists of five essays and this general introduction. While the essays are all free-standing and can be read independently of each other, there are also various connections between them. The purpose of this introduction is to summarize the essays and explain how they are related, in part by locating them in a broader philosophical context. In particular, there are at least three broad topics that relate to several of the essays. I will briefly describe them below.

The first topic concerns the idea that our normative views seem to guide us in our lives, in a way that our other views do not. For example, if I conclude that I should right now drink some coffee, then I will normally do just that. Similarly, if I find that my beliefs about certain questions are unjustified, or poorly supported by my reasons, then I will normally abandon them. There are, of course, exceptions: I can be struck by weakness of the will, for example, and for that reason fail to do what I judge that I ought to do. Or my beliefs might be so deeply ingrained in my worldview that I fail to give them up, although I realize that they are unjustified. However, even when such exceptions are taken into account, it remains a striking fact that at least for the most part, we tend to adjust our actions and attitudes so that they ‘fit’ our normative views.

Ideas along these lines are influential in the literature. For example, Frank Jackson writes that

\[\text{[it] is fine for a theory in physics to tell us about its central notions in a way which leaves it obscure how to move from those notions to action, for that passage can be left to something which is not physics; but the passage to action is the very business of ethics. (Jackson 1991, 467)}\]

\(^1\) This is simply how I use the term ‘normative’. It can also be used in other, narrower, ways.
In a similar vein, Jacob Ross also proposes that our views about what we ought to do are supposed to guide us:

…in genuine deliberation, we are guided, at least implicitly, by the question “what should [sic] I do?” or “what ought I to do?” And we ask this question not simply in order to satisfy our curiosity, but in order to make up our minds about what to do, that is, in order to form an intention. Thus, the role of the ought of practical deliberation is to guide our intentions, and thereby to guide our actions… (Ross 2012, 164)

If we find the general idea of normative guidance attractive, it raises several questions in normative theorizing. For example, how does normativity relate to other things that also guide us? In particular, many have thought that our desires can also guide our actions, in the sense that they are the kind of mental states that can motivate us to do certain things. For example, when I perform an action such as drinking coffee, I normally perform the action because I want to do it. Since both desires and normative judgments thus seem capable of guiding us, it has been suggested that the connection between normativity and desires must somehow be very intimate. While this connection can be understood in different ways, one possibility is that normative truths must in some way depend on some relevant motivational states. This idea is critically examined in Essay 1 (‘The Entanglement Problem and Idealization in Moral Philosophy’) and Essay 2 (‘Weighting Surprise Parties: Some Problems for Schroeder’). I return to those essays in section 3 of this introduction.

The second recurring topic concerns normative knowledge. It is natural to think that we have a lot of it. For example, surely we know that it is good to be nice at least sometimes, or that we should not believe everything that we are told? At the same time, however, the idea that we are capable of knowing things such as these also raises puzzles. For example, it is difficult to say exactly how we could know of them. At a minimum, many have thought, the ways in which we can acquire normative knowledge seem quite different from the ways in which we can acquire knowledge about other truths, such as those revealed to us by the sciences.

Of course, the mere fact that something is puzzling is not in itself a reason to deny that it exists—after all, the world is a puzzling place. However, the idea that we can have normative knowledge has also faced more direct challenges. In particular, the idea that we can have moral knowledge has often been viewed as especially suspicious. Two of the most influential challenges to this idea appeal, respectively, to (i) the fact that moral disagreement is both deep and widespread, and (ii) to the fact that evolutionary factors such as natural selection appear to have had a significant influence on our moral views. Essay 3 (‘From Evolutionary Theory to Moral Skepticism, via Disagreement’), which is co-authored with Folke Tersman, proposes that these challenges can be combined in a way that strengthens the case for moral skepticism. More
exactly, what we seek to show is that, at least given certain popular assumptions about the nature of moral truths (assuming that they exist), we do not have knowledge of those truths. I return to this essay in section 4 of this introduction. Essay 5 (‘Meta-Skepticism’) then turns to the more general question of whether we can know anything at all, including apparently obvious facts such as that we have hands, and proposes a new approach to it. I return to this essay below.

At the intersection of the two topics just presented—i.e., normative guidance and normative knowledge—lies a set of questions that have received a lot of attention recently: Is there a kind of normative guidance that we can hope to get even in choice situations where we do not know the answers to the relevant normative questions? Alternatively: What should we do when we are uncertain or ignorant about what we should do? And how should questions such as these even be understood? While it is both natural and tempting to think that some form of normative theorizing could guide us even in choice situations where we are normatively uncertain, it is unclear how the relevant questions should even be formulated. After all, if I do not know what I should do, then the question of what I should do is precisely the one that I do not know how to answer. So how can we make sense of the question of what I should do when I do not know what I should do?

Essay 4 (‘Ethics and the Question of What to Do’) suggests an account of the question that normative uncertainty invites, but which also applies to similarly puzzling questions that are invited by other choice situations. As an example, consider choices in the face of normative conflicts. It is plausible that what we morally ought to do sometimes clashes with what we prudentially ought to do in such a way that we cannot satisfy both ‘oughts’. For example, there may be situations in which I morally ought to promote the common good but prudentially ought to promote my own good. These situations resemble those that involve normative uncertainty in that they also invite puzzling, apparently ‘practical’ or ‘deliberative’ questions such as: which ought—the moral or the prudential one—ought I really to satisfy? Indeed, the question that is highlighted here has sometimes been called ‘the central deliberative question’ (see, e.g., Lord 2017 and McPherson 2018).

The view that I propose in Essay 4 is that, contrary to appearances, the central deliberative question is not ultimately a question of what we ought to do, in any sense of ‘ought’—not even in some privileged or ‘all things considered’ sense. Rather, the central deliberative question is the question of what to do. While the distinction between what to do and what one ought to do might seem to be merely terminological, I further suggest that a lot turns on it because these questions should be understood in radically different ways. I return to this essay in section 5 of this introduction.

The third topic concerns the fact that while our normative concepts appear to be capable of guiding us in the ways suggested above, we can also in various ways guide and evaluate our concepts. For example, if we find that the
concepts that we are employing are not suitable for certain purposes, we can choose to introduce new concepts. This is frequently done both in philosophy and in other forms of theorizing that rely on technical notions that are distinct from (though sometimes related to) the ones that we use in everyday life. Similarly, we can also for various reasons choose to abandon concepts that we have previously employed. For example, we might find that they do not correspond to anything in reality (like the concept of phlogiston), or that they are immoral (like the concepts expressed by various slurring terms), or that they are beset with both these problems at once.

The activity of guiding and evaluating our concepts in this way has become known as ‘conceptual engineering’ (see, e.g., Haslanger 2000, Burgess and Plunkett 2013, and Cappelen 2018; see also Burgess and Plunkett 2013 for the related idea of ‘conceptual ethics’). While this topic has been heavily debated recently, challenging questions remain about its philosophical foundations. In particular, special problems arise when we try to guide the concepts that normally guide us. For example, while there is no obvious problem in asking whether we ought to use the concept of phlogiston, there are obvious problems in asking: ought we to use the concept of ought itself? On the one hand, the answer might seem to be a trivial ‘yes’. On the other hand, however, one might also think that there should be some non-trivial question in the neighborhood, one to which the trivial answer does not do justice. This problem, as it arises concerning the concept of ought, is another problem to which the account developed in Essay 4 is applied. Essay 5 then turns to corresponding questions about our epistemic concepts, such as the concept of knowledge, in the context of skepticism about the external world. I return to that essay in section 6 of this introduction. Finally, in section 7, I provide short summaries (or ‘abstracts’) of each essay.

Before that, however, I will in the next section make some general distinctions and remarks about normative theorizing, in order to clarify the questions that the essays in this thesis are intended to answer.

2. The Study of Normativity: Some Distinctions

One central normative question is: what should we do? A question that may or may not be different is: what ought we to do? Whether or not these questions are in fact different, I will follow the common practice of using ‘should’ and ‘ought’ as synonyms. While most papers in this thesis primarily focus on the question of what we should (or ought to) do, I use the term ‘normative’ in such a way that other questions also count as normative, including questions about what is good and bad, right and wrong, just and unjust, virtuous and
vicious, and so on. These other questions are somewhat less central to the main aims of this thesis, however, since they are less closely connected to the three topics presented in section 1. In particular, the idea that normativity is in some sense action-guiding is arguably especially plausible with regard to the question of what we should (or ought to) do. To illustrate, suppose that you are in a situation in which it would be brave to run into a burning building to save a stranger. It seems that you might still ask for a kind of normative guidance here, since you might wonder whether this is the kind of situation in which you should be brave.

Another intuitive distinction can be drawn between different kinds of norms and values. More exactly, many agree that we can distinguish between the norms and values of morality, those of prudence, those of epistemology, those of aesthetics, and so on (though for exceptions, see, e.g., Crisp 2006 and Tännö 2010). In this thesis, I will both consider questions that pertain specifically to morality, questions about the relationship between morality and other kinds of normativity, and questions about normativity more generally. At some points, however, the distinction between different kinds of norms and values will be set aside. For example, my arguments against the idea that normative facts depend on motivational states are meant to be successful no matter what kind of norms and values we focus on.

Yet another distinction can be drawn between first-order and second-order normative questions. First-order questions include questions about what we ought to do, what is right and wrong, good and bad, and so on. They also include explanatory questions, such as: when an action is right, then what makes it right? (For further discussion see Fogal and Risberg forthcoming.) In contrast, as their label suggests, second-order normative questions can be viewed as questions about first-order normative questions. These questions include ontological questions such as: are there truths about what we ought to do, what is good, and the like? If so, what is the nature of such truths? For example, are they very different from the truths revealed to us by the natural and social sciences, or by mathematics? They also include epistemological questions such as: if there are truths about what we ought to do and what is good, then how do we figure out what those truths are? Are some methods for forming normative judgments better or more reliable than others, and if so, which ones? Are we even capable of figuring out what we ought to do, what is good, and so on? Finally, second-order normative questions include broadly psychological ones, such as: What is it to judge that something, such as an action, is good, or that it ought to be performed? Is it to form a belief about some matter of fact? Or is it more like acquiring a positive attitude or emotion towards the thing in question? These are not the only kinds of second-order

---

2 Cf. Schroeder (2005, 3), who by ‘the normative’ means “the whole class of subject matters which are of broadly the same kind as those traditionally studied by ethical theory: the right, just, good, rational, correct and so on”.
normative questions that there are; for example, such questions also include broadly semantic or linguistic ones about the meaning of terms like ‘ought’ and ‘good’. But they are the ones that will be most central in this thesis.

The study of second-order questions about ethics or morality is commonly called ‘metaethics’. Clearly, however, similar questions can also be raised about the ontology, epistemology and psychology of other normative domains, such as prudence, epistemology, and aesthetics. Second-order questions about all these domains (and the relations between them) can thus be called ‘metanormative’ questions.

This thesis mainly focuses on second-order questions about ethics and normativity. That said, many of the discussions are informed by, and have consequences, for issues in first-order normative theorizing as well. What this illustrates, I think, is that it is not always obvious how the distinction between first- and second-order normative questions should be drawn, and that there are interesting connections between these fields. Some of these connections will become apparent as this introduction proceeds.

Finally, all papers in this thesis assume the truth of a view known as ‘cognitivism’, which is a view about the psychological second-order questions mentioned above. While this view can be formulated in different ways, the core idea is that normative judgments are ‘belief-like’, ‘representational’ or ‘truth-apt’ in the same broad sense that, e.g., scientific or mathematical judgments are. Cognitivism thus does not entail that normative judgements are ever true, but merely that they can be assessed in terms of truth and falsity. (Compare: since atheism implies that judgments to the effect that God exists are always false, it also implies that such judgments can be assessed in terms of truth and falsity.) The main alternative to cognitivism is ‘non-cognitivism’, according to which normative judgements are not belief-like but ‘desire-like’, non-truth-apt states, such as states of approval and disapproval. It is my view that non-cognitivism is false about ethical judgments, and about normative judgments more generally. However, I also think that a form of non-cognitivism is true of a nearby question, one which can also be viewed as a ‘practical’ or ‘deliberative’ one: the question of what to do. I will return to this issue in section 5.

3. Normative Facts and Desires

As suggested above, one way to approach the idea of normative action-guidance is by considering the relation between normative judgments and other things that guide us, such as our motivational states, and in particular our desires.3 In the first two essays in this thesis, I critically examine views that are motivated by the idea that this relation is especially intimate. Because those

---

3 It is not the only approach, of course. For criticism, see, e.g., Parfit (1997).
essays are shorter and include less by way of background than the other essays in this thesis, this section will be somewhat longer than the other sections in this introduction.

While it is controversial how and whether normativity and motivation are related, it is often agreed that motivation and *action* are intimately related. Indeed, while we sometimes do things that we are not motivated to do, those things normally seem not to be actions properly speaking, but ‘doings’ of some other kind. Examples of such things include physiological reactions, like unintentionally sneezing, and other involuntary events, like falling down the stairs. While these are also things that we do (in a broad sense of ‘do’), they are normally neither actions nor things that we are motivated to do.\(^4\)

Moreover, many philosophers believe that motivation is in turn closely related to *desires*. This view accords with a popular theory in the philosophy of mind known as ‘the Humean theory of motivation’.\(^5\) According to this theory, an agent’s having a belief about the world is never sufficient for her to be motivated to act. Instead, according to Humeans, motivation also requires the presence of a desire (or, at least, of a ‘desire-like state’ in some broader sense). Thus, by appealing to the Humean theory, one can argue that desires are closely connected to motivation. Moreover, given the reasoning in the previous paragraph, this conclusion would also show desires to be closely connected to action.

While the Humean theory is far from universally accepted, my aim here is not to evaluate it, but merely to reconstruct a kind of reasoning that might lead one to think that desires or other motivational states are in some sense central to normativity. Many popular views about both first-order and second-order normative questions imply that this is the case. For example, many views about personal welfare (in the broad family of ‘preferentialism’) entail that how good our lives are for us depends on the extent to which our desires or preferences are satisfied.\(^6\) Further, on many views about what an agent ought to do (in various senses of ‘ought’), such facts depend on the desires that some relevantly idealized subject would have in certain circumstances. This might be a so-called ‘ideal observer’, for instance, or a hypothetical version of the agent herself. There are also views in political philosophy (in the broad tradition of social contract theory) according to which the fundamental principles of justice depend on the preferences that we would have behind a hypothetical ‘veil of ignorance’, where we are assumed to be unaware of our gender,

---

\(^4\) I do not mean to suggest that whenever we do something that we are motivated to do, that thing is an action; that view faces well-known problems having to do with deviant causal chains. The point is instead that when we do something that we are not motivated to do, then that thing is usually not an action.

\(^5\) Proponents of the Humean theory include most non-cognitivists, but also cognitivists such as Smith (1994) and Railton (1986). Opponents include Nagel (1970) and Shafer-Landau (2003).

\(^6\) This dependence claim can be made more precise in different ways, depending on how exactly preferentialism is formulated. For further discussion, see Rabinowicz and Österberg (1996).
nationality, religious views, and other factors that determine our place in society. Finally, non-cognitivists think that normative judgments simply are desire-like states.

As these examples illustrate, the broad idea that desires are central to normativity can be made more precise in two different ways, depending on whether desires are taken to be related to the contents of normative judgments or to normative judgments themselves. On the first idea, it is normative facts or truths—the things that normative judgments concern—that relate to desires in a certain way. For example, if I ought to donate some money to charity, then this view entails that this action stands in some appropriate relation to some desires. On the other, non-cognitivist idea, what relates to desires is not the things that normative judgments concern, but the judgments themselves. For example, if I judge that someone ought to donate money to charity, then this view requires that this judgment stands in some appropriate relation to my desires. The action that my judgment concerns need not do so, however, and the person that my judgment concerns need not be me. For now, I will set non-cognitivism aside and focus on the first idea. More precisely, I will argue that several popular versions of it should be rejected.

There are different ways in which normative facts and desires could be related. First, as the examples above illustrate, one could focus on different kinds of normative facts; i.e., facts about welfare, justice, different kinds of oughts, and so on. For ease of exposition, I will here focus on oughts. (The distinction between moral oughts, prudential oughts and so will not be crucial in what follows.) Second, one could focus on either actual or hypothetical desires. While the latter view has been more popular, it will be helpful to begin by considering some versions of the former.

The most straightforward formulations of the idea that what we ought to do depends on our desires are obviously unsatisfactory. For example, consider the following principle:

\[(P1) \text{: An agent } A \text{ ought to perform an action } \phi \text{ if, and only if, } A \text{ desires to } \phi.\]

One of the many problems with (P1) is that, in a given situation, an agent might want to do many different things. For example, in a situation where I want to buy a cake, I might also want to buy something healthier, such as carrots. Moreover, it might be that I cannot perform both these actions (for example, I might not have enough money to afford both carrots and cake). In this situation, (P1) implausibly implies that I face a kind of normative dilemma. More precisely, it implies that I ought to perform two different actions such that I cannot perform both of them. While such dilemmas are perhaps possible, as some philosophers believe, this situation is clearly not one of them. In particular, we might add to the case the further assumption that my desire for carrots is much stronger than my desire for cake. Given this
assumption, it is plausible, even from the point of view of a desire-based theory, that the only thing that I ought to do is to buy carrots.

The standard solution to this problem is to insist that what an agent ought to do depends on her strongest desire, as follows:

\[(P2): \text{An agent } A \text{ ought to } \phi \text{ if, and only if, } A \text{ desires most strongly to } \phi.\]

‘Most strongly’ is here naturally understood in relation to the actions that \(A\) can perform in the situation. Thus understood, whether \(A\) ought to \(\phi\) depends on this view on whether \(\phi\) is the action of those that she can perform that she most strongly desires to perform. The most influential desire-based view on which the strength of our desires do not determine what we ought to do is the one endorsed by Mark Schroeder (2007). I will consider Schroeder’s view in section 3.2.7

Although (P2) avoids the problem with (P1) just mentioned, it faces many other ones. One influential objection to principles like it is due to Bernard Williams (1979) and goes roughly as follows. Suppose that an agent is thirsty and wants to drink a glass of water. There is a glass in front of her which, she believes, contains water. Thus, what she most strongly wants to do is to drink what is in the glass. However, suppose that the agent is mistaken about the contents of the glass: while she believes that it contains water, it in fact contains petrol. In this case, the argument goes, the agent ought not to drink what is in the glass, contrary to what (P2) entails.

Arguments along these lines have often been taken to refute principles like (P2) and instead motivate an appeal to idealized desires (as opposed to actual ones). However, it is less often noted that such arguments rely on controversial claims about how and whether ignorance can have an impact on what we ought to do. More precisely, the argument just presented assumes a kind of objectivism about these questions, according to which what an agent ought to do does not depend on what she believes (or has reason to believe) that the world is like, but only on what the world is in fact like.8 Accordingly, if one rejects such objectivism, one could also resist the conclusion that the agent in Williams’ case ought not to drink what is in the glass. For on non-objectivist views, the fact that the glass in fact contains petrol makes no difference to what the agent ought to do, if she is unaware (or has no reason to think) that this is the case.

---

7 Another non-standard desire-based view has been developed by Manne (2016). On this interesting ‘democratic’ view, a desire can be a reason for any agent (and not just for the agent to whom the desire belongs) to act. While I lack the space to discuss Manne’s view in more detail, I suspect that many of the problems for (P2) beset it too, since Manne holds that the weight of a normative reason depends on the strength of the desire from which it is generated (together with the degree to which the action in question ‘promotes’ the desire).

8 It is an interesting but controversial question whether this form of objectivism is compatible with the idea that normativity is supposed to be action-guiding; for further discussion, see, e.g., Way and Whiting (2017) and Essay 4 in this thesis.
The question of how uncertainty and ignorance can affect what we ought to do is central in Essay 4, and I will not spend more time on it here. Instead, at this point, it suffices to note that the appeal to idealized desires can be motivated in other ways. For example, Harry Frankfurt discusses a case of a heroin addict whose addiction is so strong that he most strongly wants to take the drug. The addict, Frankfurt writes,

…hates his addiction and always struggles desperately, although to no avail, against its thrust. He tries everything that he thinks might enable him to overcome his desires for the drug. But these desires are too powerful for him to withstand, and invariably, in the end, they conquer him. He is an unwilling addict, helplessly violated by his own desires. (Frankfurt 1971, 12)

Contrary to (P2), taking the drug is intuitively not what the addict ought to do. This problem too could be avoided, one might think, if we do not focus on addict’s actual desires, but on the ones that he would have had in some suitable hypothetical circumstances (see, e.g., Smith 1994).

The idea that what we ought to do depends on idealized desires is the target view in Essay 1, which is called ‘The Entanglement Problem and Idealization in Moral Philosophy’. While different versions of these views have been proposed, they can be schematically formulated as follows:

(P3): An agent $A$ ought to $\phi$ if, and only if, $A$ would in some suitable circumstances desire most strongly that she $\phi$s.

Before proceeding, some clarificatory remarks are in order.

First, as (P3) is a bi-conditional, it says that facts about what we ought to do co-vary in a certain way with facts about idealized desires. However, proponents of views of this type typically mean to make some stronger claim as well. First, their view is typically meant to be a necessary truth of some kind. Moreover, if the view is meant to answer first-order normative questions, one might want to add that facts about what an agent ought to do obtain in virtue of facts about idealized desires. If the view is instead meant to answer ontological second-order questions (that is, again, questions about the existence and nature of normative facts), one might instead want to add that normative facts or properties in some sense consist in or are reducible to facts or properties that have to do with idealized desires. Alternatively, if the view is meant to answer certain psychological second-order questions, one might instead add that normative concepts can be analyzed in terms of idealized desires. Although the distinction between these kinds of claims is sometimes crucial, it can in the current context be set aside. The reason is that all these claims entail necessarily true bi-conditionals of the form of (P3), and it is the plausibility of such claims of necessary co-variance that I will focus on here.
Second, notice that there is an ambiguity in the consequent of (P3). According to what is known as the example model, what matters is what the idealized agent would herself want to do in the relevant situation. According to the advice model, in contrast, what matters is instead what the idealized agent would want the actual agent, non-ideal as she is, to do. It is commonly held that the example model faces counter-examples that the advice model avoids (see, e.g., Smith 1995 and Markovits 2014, 39), so I will focus on the advice model here. However, with minimal adjustments, the argument in Essay 1 can be formulated to target either version of the view.

Third, different views can be held about how the notion of ‘suitable circumstances’ should be spelled out. At a minimum, adherents of idealization views have often taken it to involve full empirical (or non-normative) information. The idea is that the desires that are relevant to what A ought to do are ones that she would have if she were fully informed about facts about, for example, the consequences of her alternatives. This plausibly deals with the case presented by Williams: if what matters is not what the agent wants (i.e., to drink what is in the glass) but what she would want if she were informed about what is in the glass, it is easier to reach the result that she ought not to drink what is in it. The appeal to full information is thus congenial with the assumption of objectivism mentioned earlier.

However, the appeal to full information does not obviously suffice for dealing with the case presented by Frankfurt. The reason is that it seems that the heroin addict could be entirely aware of the relevant empirical facts, concerning, e.g., the long-term consequences of taking the drug, but still remain ‘conquered’ by his desires in such a way that he most strongly wants to take it anyway. Since it remains plausible that he ought not to take the drug in this case, one might want to add further assumptions about what the idealized conditions should involve. Here one suggestion is that in the suitable circumstances, the agent would also satisfy some constraints of coherence or structural rationality (see, e.g., Smith 1994). While it is unclear whether this move ultimately yields a more plausible verdict on Frankfurt’s case, the argument in Essay 1 does not depend on these details. More generally, what the argument seeks to show is that no matter how these various details of idealization theories are spelled out, they will unavoidably be extensionally inadequate.

3.1 The Entanglement Problem

To begin with, let us adopt the convention of letting ‘A+’ denote A’s relevantly idealized counterpart (i.e., a counterpart of A that is idealized in the way specified by the target theory). Consider now the following case:
The apple case: An apple lies in front of agent $A$, who enjoys apples and can eat it with no difficulty. Her only alternative to eating it is to abstain from doing so. A demon threatens to hurt $A$ badly if, and only if, $A$ eats the apple and a suitably idealized agent, $A+$, would want $A$ to eat the apple. More precisely, there are two conditions which are individually necessary and jointly sufficient for the demon to hurt $A$: that $A$ eats the apple, and that $A+$ would want $A$ to eat the apple. If one or both conditions are not fulfilled, the demon does nothing at all.

The possible outcomes in this case can be represented as follows:

<table>
<thead>
<tr>
<th>The apple case</th>
<th>$A$ eats the apple.</th>
<th>$A$ does not eat the apple.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A+$ would want $A$ to eat the apple.</td>
<td>Pleasure &amp; Pain: $A$ is mildly pleased and then badly hurt.</td>
<td>Nothing: Nothing happens to $A$.</td>
</tr>
<tr>
<td>It is not the case that $A+$ would want $A$ to eat the apple.</td>
<td>Only Pleasure: $A$ is mildly pleased.</td>
<td>Nothing: Nothing happens to $A$.</td>
</tr>
</tbody>
</table>

As this representation shows, which outcome is actualized depends both on what $A$ does and on what $A+$ would want $A$ to do. The argument seeks to show that as a consequence, what $A$ ought to do is bound to be whatever $A+$ would not want $A$ to do. All principles of the form of (P3) are incompatible with this conclusion.

To see the problem, assume first that $A+$ would want $A$ to eat the apple. Given this assumption, $A$ can either realize Pleasure & Pain by eating the apple, or Nothing by not eating the apple. While $A+$ would want $A$ to eat the apple, this is evidently not what $A$ ought to do. On the contrary, $A$ ought to abstain from eating the apple, thus ensuring that Nothing rather than Pleasure & Pain obtains. For, clearly, in a choice between nothing at all and mild pleasure followed by horrible suffering, one ought to choose the former.

The problem is not avoided if we instead assume that $A+$ would not want $A$ to eat the apple. Given this assumption, $A$ can either realize Only Pleasure by eating the apple, or Nothing by not eating the apple. While $A+$ would not want $A$ to eat the apple, it is evident that in this situation, eating the apple is what $A$ ought to do. For the choice is now between nothing at all and mild pleasure that is not followed by horrible suffering. And in this situation, one clearly ought to choose the latter.

The upshot is that $A$ ought to eat the apple if, and only if, $A+$ would not want $A$ to eat the apple. This upshot is incompatible with the view that $A$ ought to $\varphi$ if and only if her relevantly idealized counterpart $A+$ would want $A$ to $\varphi$. 

22
Notably, this argument avoids many complications in the debate about idealization views. For example, as already mentioned, it allows us to set aside questions about how exactly the idealized circumstances should be spelled out, since it does not rest on such assumptions. It also sidesteps the question of whether principles like (P3) have been sufficiently supported, since its aim is not merely to show that such principles are unjustified; rather, its aim is to show that such principles are false. The argument depends only on two first-order normative assumptions that seem to be irresistible: (a) that given a choice between Nothing and Pleasure & Pain, A ought to choose the former; and (b) that given a choice between Nothing and Only Pleasure, A ought to choose the latter. In the essay, I argue that this problem is generalizable to all theories that appeal to idealized attitudes to explain what we ought to do—attitudes that need not be either desires or ones that the agent herself would have—such as the ‘Ideal Observer Theory’ held by Roderick Firth (1952) and various related views.

In discussion, several people have expressed doubts to me about the coherence of the apple case. I am not sure, however, that I have fully understood the grounds for such doubts. One argument that is meant to show that the case is incoherent runs roughly as follows. According to the argument, we can plausibly assume that a counterpart of A deserves the name ‘A+’ only if that counterpart is benevolent (in some sense). Given that A+ would have to be benevolent, it follows (i) that A+ would not want A to do something that would cause A great pain, at least insofar as one of the actions available to A would not have that consequence. It also follows (ii) that A+ would want A to do something which would cause A pleasure, at least if this would not also lead to pain, and A’s only other option would lead neither to pleasure nor to pain. However, the argument continues, the apple case is incompatible with (i) and (ii). If A+ would want A to eat the apple, then doing so would lead to serious pain, contrary to (i). And if A+ would not want A to eat the apple, then doing so would lead to some pleasure and no pain, contrary to (ii). Thus, the argument concludes, the assumption that the apple case is possible implies a contradiction.

While it is disputable whether the assumption that A+ would have to be benevolent actually supports (i) and (ii), the argument just given is unsuccessful even if we grant that (i) and (ii) are both true. The reason is that (i) and (ii) do not imply that the case is incoherent. All that they imply is that, in relation to the apple case, none of A’s counterparts deserves the name ‘A+’. It is true that if we assume that one of A’s counterparts will deserve that name and be such that both (i) and (ii) are true, then it seems that a contradiction follows. That is a non-trivial assumption, however, and nothing forces us to make it. (After all, there are plenty of conditions which none of our counterparts satisfy; for example, none of them are taller than themselves.) Indeed, the idea that such an assumption is true can be said to be exactly what the argument seeks to question. I thus conclude that this attempt to argue that the apple case is incoherent is unsuccessful.
If none of $A$’s counterparts deserve the name ‘$A+$’, then what does (P3) entail that $A$ ought to do? At least two different answers to this question can be motivated. However, neither answer renders (P3) plausible.

First, a natural thought is that if there is nothing that $A+$ would want $A$ to do (because none of $A$’s counterparts is $A+$), then, there is also nothing that $A$, according to (P3), ought to do. However, the implication that there is nothing that $A$ ought to do in this case is implausible. If there is nothing that $A+$ would want $A$ to do, then it follows that is not the case that $A+$ would want $A$ to eat the apple. When this holds, the available outcomes for $A$ are *Only Pleasure* and *Nothing*. And given these options, $A$ ought for familiar reasons to choose the former, contrary to what (P3), on this suggestion, entails.

An alternative approach focuses on the fact that if none of $A$’s counterparts deserves the name ‘$A+$’, then the counterfactual condition in the consequent of (P3) is impossible to satisfy. More precisely, on this approach, the counterfactual sentence ‘if $A$ were in suitable circumstances, she would most strongly desire that she $\phi$’ has an impossible antecedent. Since the standard semantics for counterfactuals with impossible antecedents (or ‘counterpossibles’) yields that all such sentences are vacuously true, this motivates the idea that on (P3), it is true of all alternative actions that $A$ ought to perform them. This implication, however, is implausible as well. If it is (vacuously) true that $A+$ would want $A$ to eat the apple, then the available outcomes for $A$ are *Pleasure & Pain* and *Nothing*. On the suggestion at hand, (P3) entails that it is true both that $A$ ought to realize *Pleasure & Pain*, and that she ought to realize *Nothing*. And not only is it implausible that this situation is in this sense a normative dilemma for $A$—it is also implausible in itself that she ought to realize *Pleasure & Pain*. Instead, clearly, the only thing that $A$ ought to do is to realize *Nothing*.

### 3.2 Hypotheticalism

The argument just presented seeks to show that no theory about what we ought to do of the form of (P3) will be successful. However, the non-idealized principles discussed above—(P1) and (P2)—also faced what seemed to be serious problems. Are there any other options for a desire-based theorist?

As already mentioned, the most influential desire-based theory about what we ought to do that is not susceptible to the problems already brought up is Mark Schroeder’s so-called Hypotheticalism (Schroeder 2007). This theory is the focus of Essay 2, which is entitled ‘Weighting Surprise Parties: Some Problems for Schroeder’. Schroeder rejects the view that the strength of an agent’s desires determines what she ought to do, which has standardly been held among desire-based theorists. Instead, he endorses a more complex view on which an agent’s desires generate normative reasons for her to act in certain ways, and where what an agent ought to do depends on the weight of her reasons, but where the weight of a normative reason does not in turn depend on the strength of the desire that generated it.
According to Schroeder, the following principle captures all facts about what normative reasons we have:

**Hypotheticalism**: $R$ is a reason for $A$ to do $\varphi$ if, and only if, there is some $p$ such that $A$ has a desire whose object is $p$, and the truth of $R$ is part of what explains why $A$’s doing $\varphi$ promotes $p$.\(^9\) (Schroeder 2007, 59)

To say that an action ‘promotes’ a certain proposition is, for Schroeder, to say that the action makes it more likely that the proposition is true, compared to if the agent were to do nothing. While there are possible problems with this understanding of promotion, I will not spend more time on them here (see instead Sharadin and Dellsén 2019 for further discussion).

To see the implications of Hypotheticalism, consider the following example due to Schroeder. Ronnie, who wants to dance, is invited to a party where there will be dancing. Intuitively, Ronnie has a normative reason to go to the party. And Hypotheticalism accommodates this intuition: on this view, the fact that there will be dancing at the party is a reason for Ronnie to go to the party, since that fact is part of what explains why Ronnie’s going to the party makes it more probable that he dances, compared to if he were to do nothing.

However, to say that Ronnie has a reason to go to the party is not yet to say that Ronnie *ought* to go to the party. Since it is also plausible that this is what he ought to do (at least given many natural ways of spelling out the case), Hypotheticalism needs to be supplemented with an account of the relationship between oughts and reasons. Here Schroeder endorses the popular view that what an agent ought to do is explained by facts about the relative *weight* of the normative reasons that speak for and against her alternative actions in the situation.\(^10\) However, while the traditional view among desire-based theorists has been that the weight of a normative reason depends on the strength of the desire that generated it, Schroeder claims that many problems for desire-based theories can be avoided if that view is abandoned. Instead, Schroeder offers the following theory of the weight of reasons:

**Weight**: A reason $R$ is weightier than a reason $R^*$ if and only if it is correct to place more weight on $R$ than on $R^*$ in deliberation. (2007, 140; notation adjusted)

This theory centrally relies on the notion of ‘placing weight’ on a reason. It is not obvious how this notion should be understood. Schroeder appears to

---

\(^{9}\) Schroeder calls this claim ‘Reason’. I have adjusted the notation and, for reasons given above, formulated it as a bi-conditional rather than as the reductive thesis that Schroeder endorses.

\(^{10}\) While this view is popular, it is not universally held, and I am disposed to deny it myself (see Fogal and Risberg manuscript, where it is argued that the reasons-relation is explicable in terms of other normative notions rather than the other way around). That said, the problem for Schroeder’s position is easiest to see if we take this view about reasons and ought on board.
assume that placing weight is a kind of psychological activity, but he does not
say much about how this activity should be understood, noting only that
‘[w]hen agents deliberate, various factors come to light, and the agent places
more weight on some of these factors than on others’ (2007, 129). While that
much is of course true, note that a supporter of Weight cannot say that placing
weight on a reason is equivalent to \textit{believing the reason to have a certain
weight}, on pain of circularity. Thus, perhaps placing (some degree of) weight
on a reason \(R\) is rather something like treating \(R\) as important (to that degree)
in practical deliberation, where ‘treating’ is understood in a way that does not
involve belief. While I am not sure how to best make sense of this notion, the
argument that I will discuss does not hinge on these details.

In my view, Schroeder’s view is no less problematic than the idealizing
accounts considered above. To see its problems, we can consider a case that
Schroeder himself considers (but without noting that his view struggles with
it). Nate hates most parties, but not all of them—he thoroughly enjoys suc-
cessful surprise parties held in his honor. Today there is a surprise party wait-
ing for him at his home. Nate is currently not aware that this is so. Schroeder
holds, and Hypotheticalism implies, that ‘given Nate’s situation, the fact that
there is a surprise party waiting for him now at home is a reason for him to go
home’ (2007, 33). Let us call this fact—i.e., the fact that there is a surprise
party waiting for Nate now at home—‘\(R\)’.

The problem for Schroeder’s view stems from the fact that \(R\) is a reason for
Nate to go home if, but only if, Nate is not aware of \(R\). If Nate were to learn
that there is a party waiting for him at home—a party that would then merely
be \textit{supposed} to be a surprise party—then this would ruin the surprise. Hence,
if Nate knew of \(R\), he would hate going to the party. Thus, \(R\) is a reason for
Nate to go home only insofar as he does not believe that \(R\) obtains.

Now, since \(R\) is a reason for Nate to go home, Weight entails that it is cor-
rect for Nate to place some (positive) weight on this fact in deliberation. How-
ever, this implication simply seems to be false. Nate can only place weight on
\(R\) if he believes that \(R\) obtains. But if he were to have that belief, it would not
be correct for him to place weight on \(R\), since \(R\) would then no longer be a
reason for him to go home. Thus, contrary to Weight, \(R\) is a reason which it
would not be correct to place weight on.

Krister Bykvist has pointed out to me that the argument just given rests on
an assumption that is not made explicit in the essay. The assumption is that
facts about correctness are not in a certain sense ‘modally evasive’. It can be
stated as follows:

\footnote{For further discussion of this and similar cases, see also Way and Whiting (2016).}
Correctness Is Not Modally Evasive (CINME):
If it is correct for an agent A to φ, then, if A were to φ, it would be correct for A to φ.

CINME follows from (but does not entail) a more general principle of normative invariance, according to which the normative status of an action cannot counterfactually depend on whether the action is performed. (This principle is put forward by Carlson 1995, chapter 6, section 6.3, and discussed by Bykvist 2007; for relevant discussion see also Rabinowicz 1985.) The argument given above requires an assumption along these lines since, if CINME is false, a supporter of Weight can hold that although it is correct for Nate to place weight on R, it would not be correct for him to do it if he were to do it.

That said, defending Weight by rejecting CINME seems to be a desperate maneuver. Note first that the only counterexamples to CINME that stand a chance at succeeding are ones that involve unperformed actions. (If I perform φ and φ is correct, we must surely accept that if I were to perform φ, as I in fact did, then φ would be correct.12) But what does it mean to say of an unperformed action that it is correct? If by ‘an action’ we mean a concrete event that takes place in space and time, then only performed actions exist. Talk of unperformed actions must thus be seen as a convenient shorthand for something else. But what? A natural suggestion is that it is shorthand for what would have been the case if the relevant action had been performed. Given this approach, however, CINME is not only very plausible but trivially true for unperformed actions. For what it then says is that if a certain counterfactual about an unperformed action is true, then that very same counterfactual is true.

There is an alternative view of unperformed actions, however, on which such actions are abstract objects that in fact exist. On one version of this view, unperformed actions are ordered triples of agents, action-types, and times.13 (Of course, this view can also be held of performed actions.) To illustrate, suppose that I did not have coffee this morning. Then, on this view, my unperformed action of having coffee this morning is (identical to) the ordered triple <Olle Risberg, having coffee, this morning>. Given that all constituents of that triple exist, there is nothing that prevents the triple (and thus the action) itself from existing, even if the action in question is not performed.

On this view of unperformed actions, CINME is not trivially true. It is not strictly inconsistent to hold that although the relevant ordered triple has the

---

12 Counterfactuals whose antecedents and consequents are both true are sometimes called ‘true-true’ counterfactuals. That all such counterfactuals are true follows from an assumption called ‘Strong Centering’ (according to which every possible world is closest to itself) together with the standard semantics for counterfactuals (see, e.g., Lewis 1973). The view that all true-true counterfactuals are true has sometimes been criticized, however, and my argument here does not rest on it (for further discussion, see, e.g., Walters 2016). Rather, what the argument rests on is the weaker and, I think, highly plausible view that all true-true counterfactuals of the form ‘if A were to perform φ, then φ would be correct’ are true.

13 For a view along these lines, see Goldman (1970).
property of correctness, it would not have had the property of correctness if I had had coffee this morning. Nonetheless, many non-trivial claims are extremely plausible, and I submit that CINME is one of them, even if unperformed actions are abstract objects. In particular, if we are tempted by the idea that normative facts are supposed to be action-guiding in some sense, then rejecting CINME seems to be a disaster. It is not particularly action-guiding to learn that, although \( \phi \) is correct, it would not be correct if you were to do it (cf. Bykvist 2007). Thus, a supporter of a desire-based theory can only reject CINME if she also abandons a crucial source of support for her kind of view.

In conclusion, then, I take the arguments just given to show that Weight is not a correct account of the weight of reasons. What this means is that even if Hypotheticalism, considered as a view of what reasons we have, is true, it still needs to be supplemented with an account of what we ought to do. And if Weight is abandoned, then Schroeder has yet to give us one.

It may be tempting to try to deal with the problems just discussed by invoking an element of idealization also in our theory of the weight of reasons. For example, consider a counterpart of Nate—‘Nate+’—who is fully empirically informed (and who thus knows about the surprise party waiting for Nate). One might think that it would be correct for Nate+ to weigh the reasons that Nate has in a certain way. In particular, the fact that Nate+ would be aware that a surprise party is waiting for Nate is no obvious obstacle to Nate+’s correctly assigning weight to that fact in deliberation about what Nate ought to do. Thus, one might think, the problem can be avoided if we modify Weight so that it concerns what would be correct for Nate+, rather than Nate, to do.\(^{14}\)

The problem with this move is that even if it deals with the surprise party problem, it instead allows the entanglement problem to arise again. Desire-based theories thus appear to face a kind of trilemma. If they involve idealization, they face the entanglement problem. If they do not involve idealization, the most natural options are either to take Schroeder’s route or to accept some rather simplistic non-idealized theory along the lines of (P1) or (P2). However, Schroeder’s view faces the surprise party problem, and theories of the form of (P1) and (P2) are, as we have seen, problematic for other reasons. While it cannot be ruled out that there are other options that a desire-based theorist could consider, of course, I nonetheless conclude that there is at least strong reason to think that what we ought to do simply does not depend on our desires in any of the ways specified by these views.

---

\(^{14}\) Thanks to Krister Bykvist for this suggestion.
4. Skepticism About What We Ought to Do

The second recurring topic in this thesis concerns normative knowledge. In the current section, I will more precisely focus on moral knowledge, since this is the topic discussed in Essay 3. The title of that essay is ‘From Evolutionary Theory to Moral Skepticism, via Disagreement’ and it is co-authored with Folke Tersman. What we seek to show is that, if there are moral facts of the kind posited by ‘moral realists’—i.e., moral facts that are independent of our attitudes, conventions, and the like—then we do not know what those facts are. Essay 4 then turns to the question of what we should do when we lack moral knowledge, and normative knowledge more generally, and considers how this question should be understood. I return to that topic in the next section of this introduction. In the current section, I will briefly present some arguments from the debate in moral epistemology and explain how they differ from the one that we offer in Essay 3.

One strategy for arguing that we have no moral knowledge is to argue that there simply are no moral truths for us to have knowledge about. That there are no moral truths is a consequence of the view known as ‘moral error theory’, which roughly states that all moral judgments are false. The error theory is primarily a metaphysical and semantic thesis rather than an epistemological one, however, and most direct arguments in favor of it accordingly have more to do with metaphysics and semantics than with epistemology. Since this is not the strategy that we pursue in Essay 3, I will here set it aside.

The metaphysical arguments in favor of moral error theory typically seek to show that the existence of moral truths requires something of the world that, according to the arguments, would be metaphysically ‘weird’ or ‘queer’. At least three different ideas about why, more exactly, such truths would be queer can be found in the literature: (i) because they would supervene upon ‘natural’ (or ‘descriptive’) facts in an inexplicable way (McPherson 2012); (ii) because anyone aware of them would necessarily be motivated to act in a certain way (Mackie 1977); and (iii) because they would entail the existence of attitude-independent demands (Olson 2014, Moberger 2018). For different reasons, however, I suspect that all of these arguments are problematic. The first one fails, I believe, because we can explain why moral facts are modally connected to natural facts in the pertinent way, by appealing to fundamental moral principles that are necessarily true (see Fogal and Risberg forthcoming, sections 2–3). The second argument is problematic because even if it would be weird if there were facts that motivated anyone who knew of them, it is hard to see why moral facts could not exist without this feature. Indeed, on the contrary, the apparent possibility of weakness of will (see sections 1 and 2) suggests that knowing what one ought to do does not entail motivation. Finally, the third argument centrally relies upon the intuition that attitude-independent demands would have to be weird and thus unacceptable to posit. The status of this intuition is contested, however, not least since many philosophers simply do not share it (for example, see Enoch 2011, 136, note 2). While different views can be had about the evidential value of intuitions that are disputed in this way, I am inclined to think that it is often rather low, in particular when the disagreeing parties are in other respects equally well equipped for figuring out the truth (for more on these issues, see Essay 3 in this thesis, Risberg and Tersman 2019, and Risberg and Tersman forthcoming). For these reasons, I am inclined to think that the epistemological arguments for skepticism are more forceful than the metaphysical ones. (Note however that there are also other kinds of arguments in favor of the error theory, such as the semantic one presented by Loeb 2008.)
Our strategy is instead to argue that moral beliefs fail to constitute knowledge because they lack some feature other than truth that knowledge requires. In other words, the idea is that moral beliefs do not constitute knowledge even if they are true. Since the supposition that moral beliefs are true is inconsistent with the error theory, this kind of skeptical conclusion has to be supported on other grounds.

It is worth noting that skeptical theses can come in different forms. For one, they may involve different epistemic concepts: a skeptic might claim that our moral beliefs are unjustified, for example, or that we ought to give them up, or that they fail to constitute knowledge even given their truth. They can also differ in their modal strength, depending on whether they are meant to be necessary truths of some kind (e.g., nomologically or metaphysically necessary ones), and in their scope, depending on whether they target all moral beliefs or only some subset of them. For the purposes of this introduction, however, I will set the differences between these versions of moral skepticism aside.

If one has argued for some skeptical view about morality (for example, that moral beliefs are unjustified), one can attempt to use that conclusion to argue for moral error theory as well. Thus, just like moral error theory can be used to support moral skepticism, skepticism may also be used to support the error theory. An argument of the latter kind would have to proceed via some principle to the effect, roughly, that if our beliefs in a certain domain are unjustified, then we should think that there are no facts in that domain. Such principles are controversial, however, and at least they do not seem to be true without exception. For example, while all beliefs about exactly how many mosquitoes there are in the world are surely unjustified, there is still a fact of the matter about what that number is (at least barring vagueness and the like).

There is more to say about these issues, including about whether the reasoning hinted at can plausibly be extended to the case of morality. Nevertheless, I will here focus on the case for moral skepticism, rather than on the further question of whether such skepticism supports the error theory, since it is the plausibility of skepticism that is most closely connected to the questions discussed later in this thesis (for further discussion of these issues, see instead Tersman 2019).

Several arguments for moral skepticism have been presented in the literature. At a minimum, in addition to the one that appeals to moral error theory, we can distinguish between four different kinds:

(a) arguments that appeal to some kind of epistemological empiricism (see, e.g., Mackie 1977, 39);
(b) arguments that appeal to some kind of causal epistemological theory (see, e.g., Korman and Locke forthcoming);
(c) so-called ‘debunking’ arguments, that support skepticism by appeal to empirical explanations (such as evolutionary ones) of why we have certain moral beliefs (see, e.g., Joyce 2006 and Street 2006); and
(d) arguments that appeal to facts about moral disagreement (see, e.g., Tersman 2006 and Tolhurst 1987).

In Essay 3, we focus on (c) and (d). What we suggest, more precisely, is that the evolutionary debunking argument and the argument from moral disagreement can be combined in a mutually supportive way. For this reason, we also propose that our argument avoids some problems that have been taken to beset other skeptical arguments, including (a) and (b). In what follows, I will say a few words about (a) and (b) before briefly outlining how our alternative argument is meant to work.

The core idea of arguments in (a) is that some kind of empiricist view in epistemology shows moral skepticism to be true. On this strategy, the epistemic problem with morality ultimately stems from the (alleged) fact that moral judgments cannot be justified by observations. The arguments in (b) instead rest on some kind of causal epistemological theory. The most simple version of this theory states that a belief that \( p \) is justified (or constitutes knowledge) only if that belief is causally explained by the fact that \( p \). This thesis is often taken to be especially problematic for adherents of a thesis known as ‘non-naturalism’, according to which moral facts differ fundamentally from the facts dealt with by the natural and social sciences, since this thesis is typically taken to suggest that moral facts are causally inert (though for an exception, see Oddie 2005).

The arguments in (a) and (b) are related, and not always clearly distinguished, because our observational beliefs are often explained by the facts that they concern. In some cases these strategies are also intertwined; for example, in an influential argument due to Gilbert Harman (1977), it is a central premise that moral facts do not figure in the best causal explanation of any observation. Nonetheless, these strategies should be kept apart, since a belief in \( p \) may be explained by the fact that \( p \) even if the belief is not observational. For example, imagine that somebody has hypnotized me to believe that I have been hypnotized. Since my belief in this case need not be observational, the kind of principle that underlies (a) suggests that it is unjustified. The kind of principle that underlies (b) does not, however, since my belief is causally explained by the fact that it concerns.

While both these arguments for moral skepticism have been influential, the epistemic principles that they rely on face problems. First, those principles

---

16 Alternatively, one might require that the belief in \( p \) is in some more general sense ‘explanatorily connected’ to the fact that \( p \) (see, e.g., Korman and Locke forthcoming). Note that while adherents of the arguments in (a) and (b) sometimes focus on justification rather than knowledge, showing that a belief is unjustified suffices for showing that it does not constitute knowledge, given that knowledge requires justification.

17 For an interesting argument to the effect that that explanatory considerations in fact favor non-naturalism—or at least a form of ‘non-reductive realism’ which has a lot in common with non-naturalism—see Roberts (2016). (Roberts prefers the label ‘non-reductive realism’ because the natural/non-natural-distinction is, in her view, problematic (Roberts 2016, note 3).)
threaten to over-generalize, as they arguably lead to skepticism about many other areas besides morality. After all, it is plausible that most other apparently a priori domains, such as philosophy and mathematics, also deal with facts that are causally inert, and with questions that cannot be settled on the basis of observations. Second, and partly as a result, those epistemic principles also threaten to be self-undermining in a certain sense. More precisely, it does not seem that we can settle by appeal to observations whether the relevant empiricist theses about epistemology are true. And it similarly does not seem that the belief that knowledge or justification requires certain causal connections itself stands in such connections to the (alleged) fact that knowledge or justification requires certain causal connections. For this reason, those principles threaten to entail that our beliefs in them are just as epistemically problematic as, according to the arguments, our moral beliefs are. This view is not strictly contradictory, of course, since it is logically possible that the truth about knowledge could not be known, and that the truth about justification could not be justifiably believed. Nonetheless, it is hard to see how a convincing case for moral skepticism could be based on ideas about which we should be equally skeptical if those ideas are true.

Whether or not these problems are ultimately devastating, they at least suggest that (a) and (b) fail to capture what, if anything, that could be distinctively epistemically problematic about moral beliefs. To put the point another way, there are all sorts of familiar epistemic worries that one might have about philosophy and other a priori disciplines, and since ethics is one of those disciplines, it is unsurprising that it too is susceptible to such worries. But there would be more bite to a form of moral skepticism which held that, even if general skepticism about the a priori can be dealt with, there is something that makes moral beliefs especially epistemically suspicious. As far as I know, skeptics have so far failed to provide a convincing argument to that effect. It is thus interesting to explore whether such an argument can be made.

In our essay, Folke Tersman and I seek to do just that. To repeat, instead of focusing on (a) and (b), our proposal is a combination of (c) and (d). While different kinds of empirical conjectures can be invoked in (c), our version focuses on evolutionary considerations and is thus a version of the so-called ‘evolutionary debunking argument’. In standard formulations of this argument, the basic idea is that natural selection and other evolutionary forces have significantly influenced our moral views. In particular, the thought goes, those forces appear to have pushed us towards embracing certain moral views independently of where the moral truths lie—all that mattered was that it was advantageous for our ancestors to embrace them. If this account of why we

\[\text{For example, Street (2006, 115) suggests that this is plausible concerning common-sensical beliefs such as that we have greater obligations to help our own children than we do to help complete strangers, that we have reason to admire and praise altruism, and that we have reason to shun or punish people who deliberately harm others.}\]
have the moral beliefs that we have is correct, the argument continues, it seems that it would be a sheer coincidence if our beliefs are true. And this conclusion in turn appears to exclude that those beliefs constitute knowledge, even if they are in fact true.

Arguments from moral disagreement, (d), appeal to different considerations to establish a similar conclusion. The point of departure of these arguments is instead the fact that disagreement about moral issues appears to be both widespread and often difficult to resolve. For example, some people are utilitarians while others are Kantians; some believe that capital punishment is just while others do not; and so on. In some cases, there might be explanations of the disagreements which are friendly to a non-skeptic. For example, it might be that one party is (to a greater extent than her opponent) influenced by bias, reasoning poorly, or overlooking relevant empirical evidence (for suggestions along these lines, see, e.g., Boyd 1988 and Sturgeon 1988). However, it is difficult to argue that all moral disagreements can be explained in this way. On the contrary, there seems to be many cases in which neither of the disagreeing parties is in a worse position than the other in relation to the aim of uncovering the moral truth. And what the argument seeks to show is that when there is disagreement between parties whose epistemic positions are in this sense equally good, it would once again be a sheer coincidence if one of them (rather than her opponent) has ended up getting things right. This conclusion, again, seems to exclude that either party has knowledge, even if one of them has a true belief.

Both (c) and (d) thus seek to show that the epistemic problem with morality is that it is in some sense a coincidence if we have reached true moral beliefs. The intuition that coincidences are epistemically problematic is well known, for example from the literature on Gettier cases (Gettier 1963; see also Faraci 2019), but it is controversial how it is best captured. In Essay 3, we propose a principle about knowledge that, we think, explains at least one central aspect of this intuition. According to this principle, a belief that \( p \) constitutes knowledge only if the belief has a modal property that is sometimes called adherence. Roughly, on our account, what this means is that the belief has been caused by the agent’s cognitive skills or abilities, or by other features of her epistemic position, in such a way that she could not easily have failed to believe that \( p \) unless her epistemic position were worse or \( p \) were false. We further propose that when evolutionary accounts of moral beliefs are combined with facts about moral disagreement, they suggest that we do not have cognitive skills of the relevant kind. The upshot of this argument is that moral beliefs are not adherent and thus do not constitute knowledge.

One notable feature of this ‘argument from adherence’ is that unlike (a) and (b), our argument at least does not immediately generalize to other a priori domains. In particular, even if mathematics and other a priori disciplines deal with questions that cannot be settled empirically and with facts that are causally inert, one may still insist that some features of our epistemic position
ensure that we can have adherent beliefs about those facts. This is not to say that our argument entails that we do have knowledge in those areas, of course. (Indeed, since the adherence requirement states only a necessary condition for knowledge, it would be a mistake to think that the argument even in principle could have this consequence.) The point is just that unlike the other arguments, ours does not immediately rule out that we can have knowledge in those domains. For this reason, we think that our argument is more likely to capture what might be distinctively epistemically problematic with moral beliefs.

5. The Question of What to Do

If the argument presented in Essay 3 is persuasive, it motivates at least some degree of uncertainty about what is good, right, and what we ought to do. For this reason, I am myself inclined to think that such uncertainty is often sensible (and to a higher degree than we typically acknowledge).

That said, even if one is not convinced by any such highly general argument for moral skepticism, one may still reasonably think that some degree of moral uncertainty—and of normative uncertainty more generally—is often motivated by other considerations. For example, some normative questions depend on difficult empirical issues. Whether it is best to adopt a certain environmental policy or to give somebody a certain medical treatment, for instance, depends on the consequences of the policy or treatment in question. Since such empirical matters may be difficult to assess, some degree of uncertainty about the normative questions that depend on them seems to be warranted as well.

Moreover, independently of empirical difficulties, some normative questions are simply very challenging in their own right (cf. MacAskill, Bykvist, and Ord forthcoming, chapter 1). To take just one example, consider the many questions of population ethics and our duties towards future generations. These questions are surrounded by problems, paradoxes, and ‘impossibility theorems’ that no theory of population ethics has avoided in an entirely satisfactory way (see, e.g., Parfit 1984 and Arrhenius 2000). Because of these difficulties, again, some degree of uncertainty about our duties towards future generations seems to be warranted.

While these considerations suggest that we should be uncertain about normative questions that are clearly challenging, they do not immediately suggest that we should be uncertain also about normative questions that seem to be easier. Other considerations do indicate this, however—indeed, they suggest that there is room for doubt even about normative questions whose answers might strike us as obvious. For example, we normally assume without hesitation that we are not acting wrongly when we brush our teeth. However, this assumption conflicts with the well-known ethical theory known as ‘maximizing act-consequentialism’ (or ‘consequentialism’ for short), which roughly states that we always ought to perform the available action with the
intrinsically best outcome. This view entails that we act wrongly whenever we could have done something better than what we actually did. And since brushing our teeth is normally not the best thing that we can do (for example, giving money to charity might often be better), consequentialism entails that our common-sensical judgments about the permissibility of brushing one’s teeth are mistaken.

Of course, one may respond by insisting that if consequentialism implies that brushing one’s teeth is normally wrong, then this is simply a counterintuitive, problematic implication of the view. (This is related to the charge that consequentialism is ‘too demanding’; see, e.g., Kagan 1984.) However, while this response is not unreasonable, it would be unreasonable to take it to show decisively that consequentialism is false. After all, even though consequentialism has some counterintuitive implications, it also has significant advantages and able defenders (see, e.g., Feldman 1986, Smart 1956, and Tännsjö 1998). Moreover, it is at any rate unclear whether the implications of consequentialism are significantly more counterintuitive than those of other comprehensive ethical theories. Thus, even if we think that consequentialism is probably false, it seems reasonable to assign at least some credence to it. And as long as we acknowledge that this view at least might be true, we should also leave some room for doubt as to whether the actions that we normally take to be permissible are in fact permissible.

The kind of reasoning just presented might not apply to each and every normative question. However, it does apply to a great many of them—sufficiently many that in most actual choice situations, some degree of uncertainty about what we ought to do seems to be warranted. If this much is accepted, several pressing questions appear to arise. How should we deal with such uncertainty? How should we go on to live our lives, that is, given that we are now uncertain even about normative questions that we might have thought to be quite easy? For example, should we let our doubts about what we ought to do have some weight in our decision-making? Or should we simply follow the recommendations of the theory that we take to be most plausible, even if it conflicts with other theories whose truth we have not ruled out?

---

19 Other than consequentialism, the two most influential approaches to normative ethics are deontology and virtue ethics. For an overview of some of the problems for deontological theories, see Alexander and Moore 2016, section 4. A problem for virtue ethics is that, on at least certain otherwise promising formulations of such theories (such as those endorsed by Kawall 2002 and Kawall 2009), they run into the entanglement problem (see section 3 of Essay 1 in this thesis). While some approaches to normative ethics, such as particularism and certain forms of Rossian pluralism, may have fewer counterintuitive implications than their rivals, this is typically because they simply have no (or at least very few) determinate implications whatsoever. Accordingly, such approaches also receive no (or at least very little) positive support from having implications that cohere with our considered ethical judgments.

20 The idea that we should always follow our ‘favorite theory’ is defended by Gustafsson and Torpman (2014). The alternative view that we should at least sometimes let our moral doubts have some weight is defended by most parties in the debate, including Sepielli (2009) and
even when we are normatively uncertain, we still have to deliberate and make difficult decisions. We would like to do so in a way that is somehow sensible and informed, even if we are uncertain about what is right and good. And since we do not have time to wait until moral philosophers have discovered the pertinent normative truths, it seems that decision-making under normative uncertainty is a topic that requires serious attention.

There is a recent lively debate about normative uncertainty that has sought to provide answers to the questions just posed (see, e.g., E. Harman 2015, MacAskill, Bykvist, and Ord forthcoming, and Sepielli 2009). However, while there is a strong intuition that there is some pressing question that normative uncertainty makes salient, it is also far from obvious what exactly this question concerns. Some writers in the debate formulate their main question along the following lines: ‘I can’t figure out what I ought to do—now what ought I to do?’ But what does this even mean?

To mention just one difficulty in interpreting this question, consider again the consequentialist view that we always ought to do what is best. As this view is usually formulated, it makes no exceptions for agents who do not know that the view is true. In other words, what this view says is that we ought to perform the best action available, whether or not we know that this is so. Thus, there is no distinctive question of ‘what we ought to do when we do not know what we ought to do’—the answer to that question is simply the same as the answer to the question of what we ought to do, full stop. And the same can be said about most standard normative theories.

At the same time, of course, an agent who asks for guidance when she is uncertain about what normative theory that is true would not be satisfied if she were told that she ought to do what the true normative theory entails that she ought to do, whether or not she knows what theory that is. After all, it is precisely what the true normative theory is that she does not know. This is not the kind of answer that she is looking for. But what this indicates is that it is unclear what kind of question she is asking in the first place.

The puzzling question that is highlighted here has sometimes been called the ‘central deliberative question’ (see, e.g., Lord 2017 and McPherson 2018). Understanding this question, and the corresponding form of uncertainty about it, is the aim of Essay 4 in this thesis, which is entitled ‘Ethics and the Question of What to Do’. In the essay, I consider cases of normative uncertainty along with several other choice situations that also invite ‘practical’ or ‘deliberative’ questions that are difficult even to formulate. While each of these choice situations has received plenty of attention in isolation, they have so far rarely been given a unified treatment, and the essay thus seeks to develop an account of the central deliberative question that applies to all situations in which it arises.

MacAskill, Bykvist, and Ord (forthcoming). See also E. Harman (2015) for an argument that moral uncertainty is normatively irrelevant.
Another type of choice situation that brings the central deliberative question to the fore involves conflicting normative requirements of different kinds. Plausibly, for instance, there are situations in which we prudentially ought to maximize our own good, but morally ought to sacrifice some of our well-being for the common good. (For further discussion of such cases, see, e.g., Baker forthcoming and Chang 2004.) In such situations we may ask: ‘Which ought—the moral or the prudential one—ought I really to satisfy?’ And it is not obvious what this might mean either—after all, it is really the case both that we morally ought to satisfy the moral ought and that we prudentially ought to satisfy the prudential ought. Because this is not the kind of answer that we were hoping to get, it is again unclear what kind of question we are asking.

At least in these respects, then, choices under normative uncertainty resemble choices in the face of conflicting normative requirements: there seems to be some pressing question that such choices make salient, but since answers to that question that seem almost trivially true are clearly unsatisfactory, it is difficult to say what, more exactly, the pressing question concerns.

Most writers on this topic have assumed that this type of uncertainty concerns a special normative question, such as what we ought to do ‘all things considered’ (cf. Lord 2017), or what we ought to do in a sense of ‘ought’ that is relevant for choices under uncertainty (see, e.g., Sepielli 2009). In Essay 4, I argue against this assumption and instead offer an alternative account that combines elements of both cognitivism and non-cognitivism (see section 2 for characterizations of these views). On this ‘divided’ metanormative view, cognitivism is true of questions about what we ought to do, but non-cognitivism is true of the practical question that sometimes remains even when all cognitive questions are settled. I thus propose that this practical question is not best understood as the question of what we ought to do, but as the question of what to do. I also suggest that this view explains the intuition that practical questions sometimes remain ‘open’ even when all factual questions have been answered.21 What has less often been noted, however, is that these questions may remain open even when all the normative questions have been answered. The divided view, I believe, best accommodates this fact.

One notable consequence of this account is thus that even if there are truths about how we all things considered ought to act—truths which might even be ‘objective’, ‘irreducible’, and so on—the central deliberative question does not concern those truths. Instead, that question cannot have a true answer. The reason is that one does not answer it by forming a belief about the world, or by forming any other kind of truth-apt mental state. Rather, I suggest, one answers this question by forming an intention to act in a certain way. (My understanding of the question of what to do thus resembles the one associated with Gibbard 2003.) On this view, then, the uncertainty highlighted above is

---

21 This intuition has been invoked by both some non-cognitivists, such as Blackburn (1998, 70) and Nowell-Smith (1954, 41), and some cognitivists, such as Korsgaard (1996).
best understood as a state of practical indecision, rather than as a state of uncertainty about what the world is like in normative respects.

An upshot of Essay 4 is thus that, in at least one central sense, normativity cannot be action-guiding. Even if we figure out what we ought to do, all things considered, this falls short of answering the question of what to do.

6. Normative and Epistemic Pluralism

The third and final recurring topic in this thesis concerns the fact that while normative concepts at least appear to be capable of guiding us, we can also guide and evaluate our concepts. In this section, I will explain how this theme shows up in both Essay 4 and Essay 5. In Essay 4, the divided metanormative account is applied to the puzzling questions raised by the thesis of normative pluralism (see Eklund 2017 and Clarke-Doane forthcoming). In Essay 5, which is entitled ‘Meta-Skepticism’, I then turn to some problems raised by the epistemological counterpart of this thesis, epistemic pluralism.

Roughly, normative pluralism is the view that the concept OUGHT is not the only ought-like concept that exists—there are also other ought-like concepts, which are distinct from the one that we use in everyday life and philosophical theorizing. While this view can be understood in different ways, one possibility is to focus on the role that the relevant concepts could play in a community of speakers. To illustrate, following Eklund (2017), let us consider an alternative community whose members use the concept OUGHT* in much the same way as we use the concept OUGHT. More exactly, while we tend to perform actions that we judge that we ought to perform, they tend to perform actions that they judge that they ought* to perform; whereas we criticize and resent people who do things that we believe ought not to be done, they criticize and resent people who do things that they believe ought* not to be done; and so on. However, suppose furthermore that OUGHT and OUGHT* are not coextensive, in the sense that some actions ought but ought* not to be performed, or vice versa. Normative pluralism can then be understood as the thesis that the situation just outlined is possible.

Whether the situation just outlined is indeed possible depends on controversial questions about normative language (such as how the reference of normative terms is determined) which it is beyond the scope of this thesis to address. That said, it is nonetheless interesting to suppose that the situation is indeed possible, if only for the sake of discussion. The reason is that given that assumption, we may again be inclined to ask questions that resemble the ones considered earlier, such as: what is so great about the ought-like concept that we happen to use? Perhaps we ought to consider using another concept, such as OUGHT*, instead? In this case too there are obvious problems in

---

22 I use small caps to denote concepts.
formulating these questions, since we are employing an ought-like concept when doing so. Thus, it might turn out that although we ought to use OUGHT, it is likewise the case that we ought* to use OUGHT*. If this is the case, there seems to be a remaining “Further Question”, as Eklund puts it (2017, 23), one in the neighborhood of whether what we ought to do or what we ought* to do is really what we ought to do. Moreover, if there are no non-trivial answers to this question—if there is simply what we ought to do and what we ought* to do and that is that—then this may seem to threaten the ‘action-guiding significance’ that we normally assign to the question of what we ought to do.23

The view of the central deliberative question suggested in Essay 4 can be applied to this problem as well. In this case, the salient ‘Further Question’ again concerns what to do rather than what we ought to do. More precisely, I suggest, it is the question of whether to do what we ought to do or what we ought* to do. Thus, as noted earlier, an advantage of the divided view is that it can in this way accommodate a wide range of cases in which the puzzling deliberative question arises. Because other views of this question have typically been informed by fewer types of choice situations, they are difficult to extend to other situations in which the question is raised. For example, while views that posit an ‘ought of uncertainty’ might explain what we naturally ask ourselves when we are normatively uncertain, they do not easily account for the questions raised by conflicting normative requirements or by normative pluralism.24 Similarly, while the idea that there is an ‘all things considered ought’ might account for the question raised by normative conflicts, it does not shed much light on the questions raised by normative uncertainty (since we might well be uncertain about what we all things considered ought to do).

Now, in analogy to normative pluralism, consider a similar sort of pluralist view about our epistemic concepts, such as our concepts of knowledge, justification, evidence, and so on. Epistemic pluralism is the view that alongside these concepts, there are many other relevantly similar concepts that we could have had instead. In Essay 5, I consider the implications of this view for some questions that concern skepticism about the external world. While skeptical theses can again differ in various ways (see section 4), I focus on the thesis that we do not know anything about what the world is like in non-mental (or ‘external’) respects, including apparently obvious facts such as that we have hands or that grass is green. That said, the problems that I will discuss can also be raised in relation to the concepts of justification, evidence, and so on.

While a variety of different skeptical challenges have been presented, the argument that I focus on in Essay 5 seeks to show that we lack knowledge about the external world because we cannot rule out that we are in fact

23 The expression ‘action-guiding significance’ is from Eklund (forthcoming). It is picked up by Clarke-Doane (forthcoming).
24 That said, the ought-of-uncertainty view also raises obvious worries about regress; see, e.g., Sepielli (2014) and sections 4–5 in Essay 4 of this thesis.
dreaming, hallucinating, or subject to some similarly distorting factor that we could not possibly detect. On this idea, for example, I cannot rule out that I am in fact a handless ‘brain in a vat’, since everything would seem the same to me if I were one. Moreover, since my belief that I have hands would be false if I were such a brain in a vat, the argument insists that I do not know whether I have hands. Finally, since there is nothing special about this example, the argument concludes that none of my other beliefs about the external world amount to knowledge either.

Anti-skeptics have sought to argue that challenges such as this one can be met. On their view, we do know that we have hands, that grass is green, and other apparently obvious facts about the external world. Thus, they typically insist that challenges like the one just outlined rely on false claims about knowledge (though for alternative anti-skeptical strategies, see, e.g., Putnam 1981 and Chalmers 2018). Skeptics have disagreed. However, both parties normally assume that the central question raised by skeptical challenges is whether we know (or justifiably believe, etc.) certain claims.

With this brief background in place, the problem that I focus on in Essay 5 can now be summarized as follows. Suppose that epistemic pluralism is true. This supposition implies that even if our concept of knowledge is ‘skepticism-unfriendly’ in the sense that anti-skeptics have tried to show, there are at least some relevantly similar, ‘knowledge-like’ concepts that are not. For example, even if knowledge (in the ordinary sense of ‘knowledge’) does not require that we can rule out that we are dreaming or hallucinating, there is some knowledge-like concept—we might call it KNOWLEDGE*—which does require this for its satisfaction. And this, in turn, might give us pause. After all, we might wonder, what is so great about the concept of knowledge that we happen to use? Might it be more important, epistemically speaking, to investigate whether our beliefs about the external world satisfy some other knowledge-like concept? And how should these questions even be understood? What I call the ‘meta-skeptical challenge’ is to make sense of the notion of epistemic importance that is naturally invoked in questions such as these. In the essay, I consider some possible ways of understanding this notion and suggest that they are all wanting.

The problem just presented is similar, both structurally and intuitively, to some of the problems considered in Essay 4 (see section 5). For this reason, it is natural to think that these problems might also have a unified solution. Unfortunately, however, there is reason to think that such a solution will not be forthcoming. In particular, while I have suggested that the central deliberative question in practical cases is the question of what to do, I do not think that this proposal can plausibly be extended to the epistemic case.25 The idea would

---

25 This is one of the ways in which my proposal differs from that of Clarke-Doane (forthcoming), who also holds that the question raised by normative pluralism is the question
be that just as the question raised in the practical case concerns what to do, the question raised by epistemic pluralism concerns what to believe—for example, whether to believe the claims that we are in a position to know, or those that we are in a position to know*, and so on. Moreover, if the analogy is to go all the way, an answer to the question of what to believe would have to be an intention to believe something, just like an answer to the question of what to do is an intention to do something.

The fundamental problem with this view, however, is that intentions to believe do not seem to be epistemically central in the same way that intentions to act are practically central. This in turn gives rise to several further difficulties. One is that the view seems to require an implausibly strong form of ‘doxastic voluntarism’, according to which we can typically intentionally control what we believe. If this view is false, as it is widely taken to be, then the question of what to believe is one that we typically cannot even seriously ask ourselves, since we can only answer it in one way. To put the point in a different way, it seems pointless to ask myself what to believe if my answer to this question cannot have an impact on what I actually believe.

Moreover, even if the truth of this strong form of voluntarism is granted, another difficulty arises instead. Given voluntarism, it seems that our answers to the question of what to believe should in many cases be influenced by considerations that are clearly epistemically irrelevant. For example, if people will be greatly harmed unless I believe that there is life on the moon, this fact should surely have an impact on how I answer the question of whether to believe that there is life on the moon (assuming now that this is a question that I can seriously ask myself in the first place). Even so, however, this fact about the consequences of having a certain belief surely says nothing about the plausibility of its content. Even though pragmatic considerations may speak for or against believing certain claims, or intending to believe them, they do not speak against the claims themselves. And this too suggests that the question of what to believe is not epistemically central in the way that the question of what to do is practically central.

I do not see it as a disadvantage that my view of the central deliberative question cannot be extended to the epistemic case. Again, while intentions to believe are not plausibly epistemically central, it is highly plausible that intentions to act are practically central, so it should be unsurprising if a correct account of central practical questions does not apply to central epistemic ones.

---

26 Of course, not everyone agrees that voluntarism is false. For example, Reisner (2013) argues that doxastic voluntarism is at least conceptually possible. Note also that this argument is compatible with thinking that we have some other form of control over our beliefs, such as what McCormick (2011) calls ‘guidance control’, or what Paul (2015) calls ‘doxastic self-control’.

27 This argument is of course related to the well-known ‘wrong kind of reason’ problem; for further discussion, see Rabinowicz and Rønnow-Rasmussen (2004). See also Reisner (2009) for a discussion of this problem in relation to the possibility of pragmatic reasons for belief.
More generally, even though the questions raised by pluralist theses about various philosophically central concepts resemble each other in certain ways, there seems to be no reason to expect in advance that a unified account of all such questions can be given. Thus, rather than assuming that such an account can be found, it seems preferable to consider each case on its own terms.

In addition to the connections between Essay 4 and Essay 5 just mentioned, there are also connections between Essay 5 and Essay 3. One has to do with a possible response to the ‘argument from adherence’ (see section 4). A moral non-skeptic may reply to this argument by holding that even if does show that there is no moral knowledge, this conclusion is in some sense not troubling, because our concept of knowledge is at any rate flawed, problematic, or unimportant. What matters, she might think, is whether our moral beliefs are justified, or whether they satisfy some alternative knowledge-like concept instead. However, this response requires that we can make sense of expressions like ‘what matters’ in epistemic contexts—and what the meta-skeptical challenge suggests is that it is unclear whether this can be done.

7. Essay Summaries

In this final section of the introduction, I provide short summaries (or ‘abstracts’) of each essay to give the reader an overview of the thesis.

*Essay 1 – The Entanglement Problem and Idealization in Moral Philosophy*

According to many popular views in normative ethics, metaethics and axiology, facts about what we ought to do or what is good for us depend on facts about the attitudes that some agent would have in some relevant idealized circumstances. This paper presents an unrecognized structural problem for such views which threatens to be devastating.

*Essay 2 – Weighting Surprise Parties: Some Problems for Schroeder*

In this article, I argue against Schroeder’s account of the weight of normative reasons. It is shown that in certain cases an agent may have reasons she cannot know about without them ceasing to be reasons, and also reasons she cannot know about at all. Both possibilities are troubling for Schroeder’s view.

---

28 For example, as Dasgupta (2018) argues, challenging questions are also raised by a pluralist view about the notion of metaphysical ‘joint-carvingness’. While the non-cognitivist view about the central deliberative question cannot naturally be extended to those questions either, this is also unproblematic, since those questions are intuitively very different as well.

29 A suggestion along these lines, though not in the context of moral knowledge specifically, is made by Papineau (2019).
Essay 3 – From Evolutionary Theory to Moral Skepticism, via Disagreement

Facts about moral disagreement and human evolution have both been said to exclude the possibility of moral knowledge, but the question of how these challenges interact has largely gone unaddressed. The paper aims to present and defend a novel version of the ‘debunking’ argument for moral skepticism that appeals to both types of considerations. This argument has several advantages compared to more familiar versions. The standard debunking strategy is to argue that evolutionary accounts of moral beliefs generate skeptical implications because they attribute those beliefs to factors that are unrelated to their truth. That strategy is vulnerable to ‘third-factor’ responses, which invoke first-order moral assumptions to question whether the Darwinian factors and the moral truths are really unrelated in that way. In contrast, our version is immune to those responses, as it does not proceed via assumptions about how Darwinian factors relate to the moral facts. Instead, it focuses on what evolutionary accounts of moral beliefs have to say about the fact that people often reach divergent moral beliefs. The argument thereby illustrates how the debunking strategy can join forces with the argument from moral disagreement. In combination, those strategies present, we think, a challenge that is more formidable than when they are considered separately.

Essay 4 – Ethics and the Question of What to Do

In this paper I present an account of a distinctive form of ‘practical’ or ‘deliberative’ uncertainty that has been central in debates in both ethics and metaethics. Most writers have assumed that such uncertainty concerns a special normative question, such as what we ought to do ‘all things considered’. I will argue against this assumption and instead endorse an alternative view of such uncertainty, which combines elements of both metaethical cognitivism and non-cognitivism. A notable consequence of this view is that even if there are objective and irreducible truths about how we all things considered ought to act, the ‘central deliberative question’, as it is sometimes called, does not concern those truths. Instead, that question does not have a true answer.

Essay 5 – Meta-Skepticism

The debate about external world skepticism has focused on whether we can have knowledge about what the external world is like. However, our concept of knowledge might be only one of many knowledge-like concepts that there are. If that is so, it follows that even if our beliefs about the external world can satisfy our concept of knowledge, there are many other relevantly similar concepts that they cannot satisfy, even if they are true. And this, in turn, might give us pause. After all, we might wonder: what is so great about the concept of knowledge that we happen to use? Might it be more important, epistemically speaking, to investigate whether our beliefs can satisfy some other relevantly similar concept instead? And how should questions such as these even be understood? In this paper, I discuss the epistemological significance of
these issues. In particular, I introduce a novel skeptical stance called ‘meta-skepticism’, which is a kind of skepticism about the idea that some knowledge-like concepts are genuinely more important than others, and suggest that it is currently unclear whether this form of skepticism can be avoided.

References


Clarke-Doane, J. (forthcoming). Morality and Mathematics, Oxford University Press. (Citations and page references are given to a draft of this book dated 23 December 2018.)


Fogal, D., and Risberg, O. (manuscript). “Reasons and Explanation (or: Reasons Last).”
MacAskill, W., Bykvist, K., and Ord, T. (forthcoming), Moral Uncertainty, Oxford University Press.
Risberg, O., and Tersman, F. (2019). “A New Route from Moral Disagree-
ment to Moral Skepticism”, Journal of the American Philosophical Asso-
ciation 5: 189–207.
Risberg, O., and Tersman, F. (forthcoming). “Disagreement, Indirect Defeat,
and Higher-Order Evidence”, in Higher-Order Evidence and Moral Epis-
and Philosophy of Mathematics”, in Explanation in Ethics and Mathemat-
Ross, J. (2012). “Rationality, Normativity, and Commitment”, in Oxford Stud-
ies in Metaethics Vol. 7, ed. Shafer-Landau, R., Oxford: Oxford University
Press.
Philosophers’ Imprint 5: 1–18.
Press.
ford University Press.
Press.
Sharadin, N., and Dellsén, F. (2019). “Promotion as Contrastive Increase in
Expected Fit”, Philosophical Studies 176: 1263–1290.
Street, S. (2006). “A Darwinian Dilemma for Realist Theories of Value”, Phil-
osophical Studies 127: 109–166.
Press.
Press.
427.
610–621.
Walters, L. (2016). “Possible World Semantics and True-True Counterfactu-