Essays on Modalized Epistemology

Jaakko Hirvelä

ACADEMIC DISSERTATION
Academic dissertation to be publicly discussed, by due permission of the Faculty of Arts at the University of Helsinki in room 12, University Main Building (Fabianinkatu 33), on the 16th of March, 2018 at 12 o’clock.

Finland 2018


Unigrafia

Helsinki 2018
Abstract

Knowledge requires more than a true belief. In order to know something one must not only believe the truth, rather one’s belief has to be related to its truth in a certain way. One of the principal aims of epistemology, the study of knowledge, is to explain what that ‘certain way’ is. This dissertation examines the idea that knowledge requires a certain modal connection to truth. Such a connection is expressed with a condition that includes modal notions, like ‘would’, ‘must’, ‘could’, ‘might’ etc. Modal notions refer not to how things are, but to how things could have been, should be, must be, or would be. Modalized epistemology is epistemology that seeks to solve epistemological problems with the help of modal notions. The motivation for modalized epistemology stems from the fact that many concepts that have a central place in epistemology seem to be modal in nature. ‘Reliability’, ‘luck’, ‘ability’ and ‘certainty’ are perhaps best explicated with the help of modalities. This thesis develops further an existing modal condition known as the safety condition and applies it to several epistemological problems. According to the safety condition in order to know it must be the case that one could not easily have erred. In Essay 1 a novel way of understanding the safety condition is offered and used to solve two perennial problems in epistemology. In Essay 2 modalized epistemology is applied to the problem of
peer disagreement. In Essay 3 it is argued that virtue-theoretic analyses of knowledge that are committed to the idea that in order to know one’s cognitive abilities have to contribute to one’s cognitive success are ultimately unsatisfactory, since the modal relation that they claim to hold between belief and its truth is too strong. The thesis contains also an introductory chapter that offers a brief historical background of the development of modalized epistemology and lays out the current debate as well as some problems that are left unanswered in the essays.
Abstrakti

suhde, joka vallitsee uskomuksen ja sen totuuden välillä tiedon tapauksissa, ei voi olla niin vahva kuin jotkut hyve-tietoteoreetikot ovat taipuvaisia väittämään. Tähän väitöskirjaan sisältyy myös johdanto, jossa luodaan lyhyt katsaus modalisoidun tieto-opin historialliseen kehitykseen sekä nykyiseen keskusteluun, ja nostetaan esiin ongelmia, joita ei käsitellä esseissä.
en för stark modal koppling mellan trosförställningar och deras sannohet. Avhandlingen har också ett kapitel som introducerar den modala kunskapsteorins historia och dagens debatt och lyftar fram några problem som inte är behandlas i de andra essäerna i avhandlingen.
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acknowledgements</td>
<td>ix</td>
</tr>
<tr>
<td>List of Original publications</td>
<td>xiv</td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Original Articles</td>
<td>41</td>
</tr>
</tbody>
</table>
Acknowledgements

The work on this doctoral thesis was supervised by Markus Lammenranta, but I am indebted to him not just for supervising this thesis, but for guiding me through my bachelor’s and master’s thesis as well. Actually, pretty much everything that I have ever done in philosophy has been supervised by Markus, and I am extremely grateful for his friendship, support, guidance and thoughtful comments throughout the years. Without him I would never have ventured into epistemology, and I am quite happy that I did. Markus is also responsible for having introduced me to Duncan Pritchard, whom he invited to visit Helsinki to give an intensive course in contemporary epistemology. Needless to say I was inspired by Duncan’s work and this fact is heavily reflected in the content of this dissertation. I have also been fortunate to receive valuable comments from Duncan at several conferences and he has been kind enough to read and comment several draft papers of mine. Duncan also hosted me at University of California, Irvine for two months, allowing me to work on new ideas in an inspiring environment. I am most grateful for his support.

I am grateful to the pre-examiners of this thesis, Adam Carter and John Greco. They provided valuable feedback that made the thesis much better and gave excellent ideas for future work. I met John in Leuven at a virtue epistemology
conference and the interest that he showed towards my work was most inspiring. I am grateful to Adam for the valuable comments he gave to me in Edinburgh.

I am also grateful to Gabriel Sandu, who has provided helpful comments on all of my previous drafts at graduate seminars and supported and encouraged me during the writing of this thesis. I am especially grateful to him for organizing and leading the European Ph.D. Net in Philosophy, which has provided me with ample opportunities to present my work and familiarize myself to the academic culture in Europe.

I am indebted to Vili Lähteenmäki, who has helped me not only by providing excellent comments on my work, but also by giving practical advice on how to survive in the academia. His collegial support has been invaluable. I would like to thank Jani Hakkarainen for inviting me to present my work at the University of Tampere and for extremely useful comments on my work. In fact two of the essays included in this dissertation have benefited from Jani’s excellent comments. I am grateful to Maria Lasonen-Aarnio for insightful comments and encouragement. Ever since I met Maria, long before I started working on my Ph.D, she has been a role model for me.

This work would look very different without the comradery of fellow Ph.D. students at Helsinki. I am most grateful to Sanna Mattila for stimulating philosophical discussions that have highlighted interesting connections between our projects and forced me to think on issues that are much more difficult than I am used to ponder. Sanna’s way of doing philosophy is extremely inspiring and I wish I had the ability to do philosophy the way she does, instead of getting continuously stuck on insignificant details. She has been a real powerhouse by
organizing numerous reading-groups which have been most useful to me. The amount of support and help that I have received from Tuukka Tanninen cannot be overestimated. His genuine interest in my work has been very inspiring and his welcoming attitude when I first started to work on my thesis made me feel at home. Co-authoring with Tuukka has been one of the greatest experiences I have had working as a graduate student. Presenting our work to an exhausted audience at the end of a long day was as close to stand-up as philosophy can be, and I am sure we gathered more laughs (not necessarily due to the content!) from the audience than a seasoned stand-up comedian would have. I am grateful to Pii Telakivi for insightful comments and friendship. Pii started a bit later than I, and she changed the atmosphere at the department to the better by organizing all kinds of communal activities that brought people together. Her work has inspired me to look at the interconnections between newest developments in philosophy of mind and epistemology, and I am thankful for her the numerous discussions we have had.

I am most grateful to all the members of the Ph.D. Net for collegial support and thought-provoking comments. Special thanks to Nils Franzén, Fabian Hundertmark, Ragnhild Jordahl, Calle Montan Elena Tassoni, Paula Tomi and Silviu Velica and Luca Zanetti. I would also like to thank the people responsible for the Ph.D. Net, including Constantin Brincus, Mircea Dumitru, Matti Eklund, Ole Hjortland, Sebastiano Moruzzi and Christian Nimtz. Their insightful comments have improved immensely the quality of my work.

I have been lucky enough to have had funding for the duration of my Ph.D. studies. I would like to thank the Academy of Finland that funded the Academy Project “The Sociality of Knowledge” I was working for during the first year of
my studies and the University of Helsinki for offering me a salaried position as a Ph.D. student. I would like to thank Heta Pyrhönen for encouragement and for directing the Ph.D. program I am enrolled in. I would like to thank Hanna Peljo for helping me with the practicalities back when I started. I thank the University of Edinburgh, Finnish Doctoral Training Network in Philosophy, University of Notre Dame and Université de Montréal for financial support.

Previous drafts of the essays this thesis is comprised of have been presented at the universities of Bologna, Bucharest Edinburgh, Helsinki, KU Leuven, Montreal, Notre Dame, Paris, Pavia, Uppsala, Tampere and Turku. I am grateful to the organizers of these events for giving me a chance to present my work and to the audiences of these events. Their comments have improved this work immensely. I am especially grateful to Tuomo Aho, Hanne Appelqvist, Natalie Ashton, Robert Audi, Bob Beddor, Fernando Broncano-Berrocal, Samuel Dishaw, Julien Dutant, Andreas Fjellstad, Daniel Fogal, Jan Forsman, Giada Fratantonio, Jie Gao, Benoit Gaultier, Sandy Goldberg, Daniel Groll, Leila Haaparanta, Raul Hakli, Ilmari Hirvonen, John Hyman, Maria Hämeen-Anttila, Liz Jackson, Carrie Jenkins, Jonathan Jenkins Ichikawa, Antti Karjalainen, Antti Keskinen, Markku Keinänen, Chris Kelp, Jaakko Kuorikoski, Jennifer Lackey, Clayton Littlejohn, Robin McKenna, Alan Millar, Ilkka Niiniluoto, Lisa Miracchi, Hichem Naar, Sara Negri, Ram Neta, Jacob Ohlhorst, Donnachadh O’Conaill, Michele Palmira, Joonas Pennanen, Tommaso Piazza, Ilkka Päättiniemi, Panu Raatikainen, Baron Reed, Arto Repo, Juho Ritola, Giovanni Rolla, Adam Sanders, Peter Schulte, Daniele Sgaravatti, Filipe Silva, Mona Simon, Declan Smithies, Ernest Sosa, Margot Strohminger, Kurt Sylvan, Tuomas Tahko, Krister Talvinen, Christine Tappolet, Folke Tersman, Teemu Toppinen, Thomas Wallgren, Tuomas Vesterinen, Jacques
Vollet, Ralph Wedgwood and Juhani Yli-Vakkuri. I would also like to thank the referee’s and editors of all the journals where I have submitted my work. While rejections never felt good they were necessary in order to improve the quality of my writing. I thank also Andreas, Markus, Sanna, Tuukka and Vili for providing detailed comments on the introductory chapter of this dissertation.

I would like to thank Niko Kärras for his friendship and support. Playing chess with Niko at Janoinen Lohi after work has been the best way to clear my head from silly philosophical ideas. I would like to thank Erkki Heino and Pyry Veteli for funny philosophical discussions and friendship. I am especially grateful to Erkki for urging me to continue doing epistemology.

I would like to thank my parents, Irma and Markku, who have been extremely supportive through the years. I doubt that I could ever have managed to complete, or even start, doing this thesis without the (hopefully healthy) trust in myself that they have bestowed upon me. I am thankful to Jouni, Noora, Leon and Halti for support and genuine interest in my work. I thank our dog Naava for long walks during which many of the ideas contained in this thesis were developed. Finally, I would like to thank my wife Kaisa, for being the most supportive, caring and loving companion I could ever have dreamed of. Her trust and understanding has meant the world to me. This work would most certainly not exist without her.

I dedicate this work to her.

Helsinki, February 2018

Jaakko Hirvelä
List of Original Publications

1. “Global Safety: How to Deal with Necessary Truths” forthcoming in *Synthese*. Reprinted with the kind permission of the publisher.


Introduction

Knowledge requires more than a true belief. In order to know something one must not only believe the truth, but one’s belief has to be related to its truth in an appropriate way. Otherwise a lucky guess would count as knowledge. One of the principal aims of epistemology, the study of knowledge, is to explain what the ‘appropriate way’ is. Before Edmund Gettier’s groundbreaking article that transformed analytic epistemology, it was presumably widely thought that a belief is knowledge just in case it was true and justified.³ Gettier demonstrated, however, that the appropriate connection cannot consist in ‘justification’, at least if justification is taken to be a fallible. To see this consider a case offered by Russell (1948, pp. 170-171), where a subject S looks at a clock that reads ‘12’ and forms the justified true belief that it is 12 o’clock. However, unbeknownst to S the clock stopped exactly 12 hours ago. S’s belief is justified because from S’s perspective she is consulting a clock that is in working order. Given the evidence that was available for S, any reasonable subject would have believed as S did. Intuitively, however, S does not know that it is 12 o’clock even though her belief is true and justified since it is only a matter of luck that the clock is showing the

³ It is debatable whether the view that knowledge is justified true belief was ever widely held among analytic epistemologists. Plantinga (1993, pp. 6-7), for example, writes “it isn’t easy to find many really explicit statements of [the justified true belief] analysis of knowledge prior to Gettier. It is almost as if a distinguished critic created a tradition in the very act of destroying it.”
correct time. Had S looked at the clock a bit earlier or a bit later her belief would have been false.

The reason why Gettier cases are possible is because justification does not entail truth. Given that a justified belief might be false, it is always possible that a belief is justified and true, while the truth of the belief is simply a matter of good luck. In essence there is a ‘gap’ between justification and truth, and that gap might be bridged either in an appropriate way or in an inappropriate way. Linda Zagzebski (1994, p. 69) has given a recipe of how to generate Gettier cases: start with a justified belief that is false due to bad luck. Add an element of good luck so that the belief turns out to be true after all. Voilà! You have a justified true belief that is true due to good luck, and intuitively fails to be knowledge.

Post-Gettier epistemology is littered with attempts to provide a fourth condition for knowledge, which together with belief, truth and justification would give the necessary and sufficient conditions for knowledge and with proposals that seek to understand justification in such a way that it is not satisfied in Gettier cases. The former attempts have generally been deemed failures, not least because they were susceptible to counter examples, as Zagzebski noted, while the latter seem to mischaracterize the concept of justification altogether.

Perhaps because of the failure to utilize the concept of justification in the analysis of knowledge, some epistemologists cleared the slate and began searching for a single condition that would capture the connection that must hold between belief and its truth in order for the belief to be knowledge. These attempts have generally been externalist in nature and have dispensed the idea that knowledge should be analysed even partially in terms of justification. A central feature of
these views is that they aim to eliminate the kind of luck that is at play in Gettier cases. The need to eliminate harmful epistemic luck gave rise to modalized epistemology.

To see this, consider Alvin Goldman’s (1967) externalist causal account of knowledge that can be interpreted as lacking a modal dimension. On Goldman’s account one knows that \( p \) if and only if the fact that \( p \) is causally connected in an appropriate way to one’s belief that \( p \) (1967, p. 369). In order to evaluate whether a subject’s belief qualifies as knowledge all we need to do is to look at how the belief is connected to its truth in the actual world. Suppose now that I see an elm in front of me and thereby come to believe that there is an elm before me. On Goldman’s account I ought to know that there is an elm before me. But assume that I cannot discriminate elms from beeches, and that right next to the elm are several beeches. Intuitively I do not know that an elm is before me because I could very easily have looked at a beech, and still believed that it is an elm. In a sense my belief is true only as a matter of good luck. What this shows is that in order to eliminate knowledge-destroying epistemic luck it is not enough that the belief is actually tied to its truth via an appropriate causal link. Rather, the belief must retain an appropriate connection to its truth even if things had been slightly different. In essence, knowledge requires a certain modal connection between belief and truth. Epistemology that seeks to explicate how knowledge must be

---

2 Whether Goldman’s account has a modal dimension or not, depends on the nature of causation. If causal conditions are modal conditions, then Goldman’s account is modal in nature. (I am indebted to John Greco for bringing this to my attention.) Be that as it may, it is easy to see that the modal condition that might be included in Goldman’s causal account of knowledge is too weak. Goldman (1976) himself seems to have been aware of this, since he retracted from the causal account and offered a modal theory of perceptual knowledge in its place.
modally related to what is known is modalized epistemology as it is conceived here.³

The first notable attempt to understand what knowledge is with the help of a modal condition was made by Fred Dretske (1971) and Robert Nozick (1981). Nozick’s formulation of the condition became much more popular, in fact so popular that quite a few seem to have forgotten that it was Dretske who first proposed the condition. The condition Nozick proposed came to be known as the ‘sensitivity’ condition. According to the sensitivity condition a subject S’s belief that \( p \) is sensitive just in case if it were the case that not \( p \), S would not believe that \( p \) (Nozick, 1981, p. 179).⁴ The condition can be presented more formally as follows with ‘\( \Box \rightarrow \)’ standing for the subjunctive conditional: \( \neg \Box \rightarrow \neg B(p) \). This condition is modal in nature because it asks to consider whether S would believe that \( p \) if \( p \) were false.

Modal conditions for knowledge are usually explicated with the help of a possible-worlds heuristic. In the possible worlds framework the space of possible worlds is centred on the actual world and branches out according to a similarity ordering. The actual world represents the world as it actually is. Possible worlds represent scenarios that could have obtained. The worlds closest to the actual world are very similar to the actual world. The worlds that are further away are less similar to it. The worlds that are on the very rim of the space of possible worlds are utterly unlike our world and perhaps even impossible to conceive for creatures like us.

³ I borrow the term ‘modalized epistemology’ from Becker (2007).
The notion of ‘similarity’ is often taken as a primitive among those who utilize a possible-worlds framework. Lewis (1973), for example, took it as a primitive notion in his semantics for counterfactuals. However, if we do not attempt to elucidate the notion of similarity beyond an ‘intuitive’ way of understanding what it is, we invite the risk of smuggling the notion of knowledge into our understanding of what makes a world more or less similar. That is, if we do not clarify what similarity is, we might end up in ranking the possible worlds in a certain order solely by virtue of our intuition of whether the subject knows or does not know in the actual world. This presents a risk for our methodology. Hence a few words on what world-similarity is are in order here.

A helpful way of conceiving world-similarity is in terms of change. World \(w\) is more similar to \(v\) than to \(u\) in case less change is needed to make \(w\) identical to \(v\) than to \(u\). The world where I spill my coffee is closer to the actual world if my coffee cup is filled to the brim than it is if my coffee cup is half full. Sometimes the modal distance between worlds can be measured simply in terms of the amount of physical change that we would need to implement to one of the worlds in order for the other one to obtain. Often we need to consider other sorts of changes as well. Suppose that I have resolved never to drive while intoxicated and that it would be against my character to act against what I have resolved to do. The possible world where I drive while intoxicated in such a case is a faraway possible world, even though the amount of (purely) physical changes that we would have to implement for it to be the case that I drive while intoxicated is roughly the same as in the case when I am a frequent drunk-driver. In the latter case, the possible world where I drive while intoxicated is much closer to the actual world than in the former case. Intentions and the quality of the changes also
matter. Some changes have greater weight than others, and we might not be able
to characterize those changes in a purely physical language. While these remarks
are rather vague, I hope they give some conception of how world-similarity is
understood here.

Epistemologists who make use of possible worlds usually assume standard
semantics for modal expressions. Thus a proposition is necessarily true if and only
if it is true in all possible worlds and necessarily false if it is false in all possible
worlds. A proposition is possibly true if true in at least one possible world and
possibly false if false in at least one possible world. A proposition is contingently
true if true in the actual world, but false in some possible world. Finally, a
sentence that features a subjunctive conditional like “if \( p \) had been the case, then \( q \)
would have been the case” is true if and only if \( q \) is true in the scope of the nearest
possible worlds where \( p \) is true.\(^5\) At this point it is worth emphasizing that
possible worlds as they are used in this context always refer to metaphysically
possible worlds – worlds that could have obtained – not to epistemically possible
worlds, worlds that are coherent from the perspective of a certain individual but
not necessarily metaphysically possible.

\[\text{Diagram}\\\]

\(^5\) Note that this interpretation of the subjunctive conditional differs from the standard interpretation
offered by Lewis (1973). For Lewis a proposition \( p \vdash q \) is true just in case \( q \) is true in the
nearest possible world where \( p \) is true. Since according to Lewis the actual world is always the
nearest possible world \( p \vdash q \) will end up being true whenever \( p \) and \( q \) are true in the actual
world. If we adopted this interpretation of the subjunctive conditional the safety condition would
end up being vacuously true whenever the subject’s belief is true. Therefore we will follow Nozick
(1981, p. 680, n. 8) and interpret the subjunctive conditional as stated above.
In this figure the actual world is denoted by ‘@’. The similarity of the worlds is presented as a distance between them. In the figure b, c, d, e, f, g and h represent worlds that are very similar to @. Worlds that are further away from @ are less similar to it, z representing the metaphysically possible world that is maximally dissimilar to @.

In the possible worlds framework the sensitivity condition reads as follows:

SENSITIVITY: A subject S’s belief that p is sensitive if and only if,

i) in the nearest possible worlds where p is false, S does not believe that p.

To evaluate whether a subject’s belief that p is sensitive we need to look at the nearest possible worlds where p is false and check whether the subject believes that p. If she does not believe that p then her belief is sensitive. At first glimpse SENSITIVITY seems like a very promising necessary condition for knowledge. It is not satisfied in Gettier cases like the stopped clock case because in the nearest possible worlds where it is not 12 o’clock S will continue to believe that it is, since the clock has stopped.

But perhaps the main reason why epistemologists have been attracted to the sensitivity condition has to do with its anti-sceptical qualities. After all, many of our mundane beliefs are sensitive. My belief that I am typing at the moment is sensitive because in the nearest possible world where I am not typing I am playing the ukulele and I do not believe that I am typing. Possible worlds where sceptical hypotheses turn out to be true – such as the possible world where I am a brain in a
vat (BIV) that is electrochemically stimulated so that it seems to me that I am typing – do not threaten the sensitivity of my belief that I am typing, since such worlds are not among the closest worlds where I am not typing. Interestingly, SENSITIVITY can also accommodate the intuition that we do not know the denials of sceptical hypotheses. This is because in the nearest possible worlds where my belief that I am not a BIV is false (e.g. in worlds where I am a BIV) everything will seem to me as it seems to me in the actual world and hence I will believe that I am not a BIV. But in that case my belief is not sensitive and I cannot know that I am not a BIV. Therefore, it seems that proponents of SENSITIVITY could eat their cake and have it too. If SENSITIVITY is a necessary and sufficient condition for knowledge then we do have knowledge of mundane matters but lack knowledge of the denials of sceptical hypotheses. It would then seem that SENSITIVITY gives the best of both worlds.

But SENSITIVITY faced at least four devastating problems. First of all, since SENSITIVITY is able to accommodate our knowledge of mundane matters, and the intuition that we cannot know the denials of sceptical hypotheses, it is incompatible with the idea that knowledge is closed under known entailment. According to the epistemic closure principle, if S knows that \( p \), and that \( p \) entails \( q \), S knows that \( q \). SENSITIVITY will violate this principle if we insert a mundane proposition in the place of \( p \) (such as I know that I have hands) and the denial of a sceptical hypothesis that is entailed by \( p \) (such as I am not a handless BIV) in place of \( q \). My belief that I have hands is sensitive because in the nearest possible world where I do not have hands I do not believe that I have hands since they have been amputated long ago. My knowledge that I have hands entails that I

---

6 This is a very rough formulation of the closure principle. See Lasonen-Aarnio (2008, p. 157) for a more accurate rendering of the principle.
am not a handless BIV. However, my belief that I am not a BIV is not sensitive, since in the nearest possible world where I am a handless BIV I continue to believe that I am not. Therefore, SENSITIVITY violates epistemic closure principles. Nozick (1981, p. 198) himself embraced this feature of his view, though most epistemologists saw the violation of closure as an adequate reason for rejecting SENSITIVITY.

Another problem that SENSITIVITY faced had to do with inductive knowledge. To use an example introduced by Sosa (1999b, pp. 145-146), suppose that you believe that the rubbish bag you just dropped down the garbage chute will soon be in the basement, on the basis of having confirmed multiple times in the past that objects that are dropped down the chute end up in the basement. Assuming that the chute is well maintained you ought to know that the bag will soon be in the basement. However, if the bag would not for some reason end up in the basement (an extremely rare event) you would still believe that it will soon be in the basement. Therefore your belief is insensitive even though it amounts to knowledge.

Thirdly, it seems that if one knows that \( p \) then it ought to be possible for one to know that one does not falsely believe that \( p \). However, the later proposition can never satisfy SENSITIVITY. After all, if it were false that one does not believe falsely that \( p \), one would believe it anyway. If SENSITIVITY were a necessary condition for knowledge, then propositions like [I do not falsely believe that \( p \)] would be unknowable, but clearly they are not.\(^7\)

\(^7\) Sosa (1999b, p. 145) makes this objection. See Vogel (2012, p. 139) for a lucid formalization of the argument.
Finally, SENSITIVITY can be applied only to cases in which the subject believes in a proposition, which is contingently true, if true at all, since the condition is vacuously satisfied if the subject believes in a necessarily true proposition. After all, if one believes that $12 \times 13 = 156$, one’s belief will trivially satisfy SENSITIVITY since there are no nearby possible worlds where it is false that $12 \times 13 = 156$. However, it is possible to believe in a necessarily true proposition without knowing it. Perhaps one used a malfunctioning calculator which by sheer chance happened to produce the correct answer as an output.

After the problems with SENSITIVITY became apparent, some epistemologists proposed a new, rather similar, modal condition for knowledge, which came to be known as the safety condition. According to the safety condition a subject S’s true belief that $p$ is safe if and only if S would believe that $p$ only if it were so that $p$. Or more formally: $B(p) \implies p$. The idea behind the safety condition is that in order to know something it must be the case that one could not easily have erred. The safety condition can be explicated with the help of a possible-worlds heuristic as follows:

**SAFETY:** A subject S’s belief that $p$ is safe if and only if,

i) in nearby possible worlds where S believes that $p$, $p$ is true.

The crucial difference between SAFETY and SENSITIVITY is that they hold different possible worlds as relevant when determining whether the subject

---

8 Nozick recognized this problem and added another condition, which he called ‘adherence’, to his analysis of knowledge. A belief that $p$ adheres to the truth just in case if $p$ were true one would believe that $p$ (Nozick, 1981, p. 179).

satisfies the condition. When determining whether a belief is safe, only the nearby possible worlds where the subject continues to believe what she believes in the actual world are relevant. SENSITIVITY, however, is not restricted to the space of nearby possible worlds. Rather, in order to evaluate whether a given belief is sensitive, we need to consider the nearest possible worlds where the proposition that the subject believes in the actual world is false. Depending on the modal nature of the belief that the subject formed in the actual world, the nearest possible worlds where what she believes is false could be a nearby or a faraway possible world. Thus, the fact that a subject could not track the truth of her belief to the possible worlds that lay at the outer rim of the possibility space could prohibit her from having knowledge. According to SAFETY, however, the faraway possible worlds are never relevant when determining whether a subject knows or not, though they might be relevant when determining whether the belief is true or not. For example, if the believed proposition contains a modal term, such as “it is possible that it is sunny tomorrow”, faraway possible worlds do count as relevant when determining whether the proposition is true. However, in order to evaluate whether the belief is safe, the belief has to match the truth only in the scope of the nearby possible worlds.

Since SAFETY is restricted to the space of nearby possible worlds, it has welcome anti-sceptical qualities. Assuming that the actual world is roughly the way we think it is, the possible worlds where sceptical hypotheses obtain are faraway possible worlds. After all, a lot would have to have to have changed if it was

---

10 But why should this assumption be granted? After all, we do not know which world is the actual world. The reason why this assumption is legitimate is because our opponent is a global sceptic who claims that knowledge is impossible. The global sceptic must allow us to pick any metaphysically possible world as the actual world and show that even if the world was just like we thought it would be, we could not have any knowledge of it. The dialectic situation would be different if we attempted to show that we actually do have much knowledge.
the case that I am a BIV, given that at the moment it is not even possible to produce BIVs that have genuine experiences. Therefore, in all the nearby possible worlds where I believe that I am not a BIV my belief is true, and hence safe. Furthermore, it is easy to see that my mundane beliefs, such as I have hands, are safe, since in all the nearby possible worlds where I continue to believe that I have hands I do have hands.

Moreover, SAFETY is compatible with epistemic closure principles.\(^\text{11}\) After all, if my beliefs that \(p\) and that \(p\) entails \(q\) are safe, then \(p\) must be true in nearby possible worlds where I believe that \(p\), and \(q\) must also be true in those possible worlds since if my belief in \(p\) entails \(q\) is safe, \(p\) will entail \(q\) in nearby possible worlds where I believe that it does. Therefore, in nearby possible worlds where I believe safely that \(p\), and that \(p\) entails \(q\), my belief in \(q\) will also be safe.

Like SENSITIVITY, SAFETY is able to deal with Gettier cases. Consider, for example, the stopped clock case. In most nearby possible worlds where S continues to believe that it is 12 o’clock her belief turns out to be false, since she looks at the clock a bit earlier or a bit later. Therefore S’s belief is not safe and does not qualify as knowledge.

Finally, though both SENSITIVITY and SAFETY conditions have been advanced as conditions that seek to eliminate knowledge-destroying luck,\(^\text{12}\) only SAFETY

\(^{11}\) Even though it is generally accepted that SAFETY is compatible with epistemic closure principles, Cohen (2004) and Murphy (2005) have argued that SAFETY violates closure principles. The version of the safety condition that is put forward in Essays 1 and 2 is not susceptible to the kind of counterexample introduced by Cohen and Murphy. Lasonen-Aarnio (2008) has argued that ‘weak’ formulations of the safety condition violate multi-premise closure if the method of deduction is understood as a fallible method. I have argued (forthcoming) that weak formulations of the safety condition do not violate multi-premise closure once they are relativized to the method of belief formation that the subject uses in the actual world.

\(^{12}\) Engel (1992) labels the knowledge-destroying luck “veritic luck”.

12
is motivated by a fully-fledged analysis of epistemic luck.\textsuperscript{13} In order to understand the ramifications of this point we need to consider what luck is and in what way it is incompatible with knowledge.

The idea that knowledge is incompatible with a certain kind of luck was already noted by Plato in \textit{Theaetetus}, where Socrates convinces Theaetetus that true belief is not sufficient for knowledge since the belief might be true as a matter of luck. In the wake of the Gettier debate, Peter Unger developed an account of knowledge in terms of luck (or accident as he called it). According to Unger (1968, p. 161), a subject knows that $p$ if and only if it is not an accident that the subject believes the truth with respect to $p$. But clearly all luck is not epistemically harmful. For instance, many scientific discoveries were the result of an accident, yet we do not think any less of them on account of this. Unger noted this, and listed three different ways in which luck is compatible with knowledge. Very briefly the three varieties of luck that are compatible with knowledge possession that Unger noted are the following: (i) you can witness an event which is itself an accident and yet know it (content luck); (ii) it can be a matter of luck that you are in possession of evidence that allows you to gain knowledge (evidential luck); (iii) it can be a matter of luck that you formed a belief in your predicament (doxastic luck).

However, he left the notion of luck undefined and took it as a primitive concept.\textsuperscript{14} As a consequence, almost no one accepted Unger’s analysis of knowledge, even though it might in fact have been counterexample free.

\textsuperscript{13} When discussing the early predecessor of the sensitivity condition, Dretske (1971, p. 18) notes that “I want to claim that any S satisfying conditions (A)-(C) is a person of whom it is true to say that it is not at all accidental that he is right about P’s being the case.” And here is Luper-Foy (1984, p. 34) when he first introduced what came to be known as the safety condition: “A belief that S has not inferred from another is nonaccidentally correct and thus noninferentially known if and only if a causal chain that meets certain reliability conditions joins it to an event or state of affairs that guarantees its truth. That is my analysis of noninferential knowledge in a nutshell.”

There are two reasons as to why epistemologists did not subscribe to Unger’s analysis. First, as Goldman (1976, p. 773) notes, Unger’s analysis is hardly satisfying “for the notion of "non-accidentiality" itself needs explication”. Secondly, as is stressed by Nimtz (2013), given that one of the central goals of epistemology is to provide an answer to the question “what distinguishes knowledge from a luckily true belief?” an account that says “well, that it is not true as a matter of luck” seems unacceptable. Therefore, it is not surprising that epistemologists have tried to come up with detailed accounts of luck and conditions that seek to eliminate the harmful epistemic luck.

There are currently two accounts of luck on the table that have notable support: the lack of control account and the modal account. Generally, the advocates of the lack of control account and the supporters of the modal account have agreed that only events that are significant to a subject can be treated as lucky or unlucky (Broncano-Berrocal, 2015; Coffman, 2007; Pritchard, 2005; Riggs, 2007), although Pritchard (2015) has in his most recent work rejected the necessity of a significance condition on lucky events. The distinction between good and bad luck is made in terms of positive and negative significance.

On the lack of control account, an event is lucky for a subject if the occurrence of the event is beyond the subject’s control. The idea that we are lucky with respect to events that are beyond our control is very intuitive. A person who wins a fair lottery is extremely lucky to have won, but if we assume that she has rigged the

15 There are also ‘epistemic’ accounts of knowledge undermining luck that invoke epistemic notions other than belief on the consequent-side of the bi-condition. Given that the purpose of anti-luck epistemology is to elucidate the nature of knowledge in terms of epistemically harmful luck, these accounts are less suitable for our purposes, even though they might be extensionally correct. For epistemic accounts of luck, see Steglich-Petersen (2010) and Goldberg (2015).

lottery so that she is in total control of which ticket will win her snatching the prize seems hardly a matter of luck. The idea that lucky events are beyond our control is also supported by the idea that responsibility and luck exclude each other. If a subject is not in control of her actions then she is not fully responsible for them either and cannot be blamed or praised for those actions. Given that we should not blame or praise agents who fail or succeed in their endeavours solely as a matter of luck, there would seem to be a tight connection between luck and lack of control.

However, there is good reason to doubt that a lack of control condition could be a sufficient condition for an event occurring as a matter of luck. As Latus (2000, p. 167) neatly puts it, the rising of the sun is beyond our control, yet we do not conclude that it was a matter of luck that the sun rose. Pritchard (2005, p. 127) points out that the lack of control condition is particularly problematic when it comes to epistemic luck, since most of our perceptual beliefs are not within our immediate control. It would be odd to conclude that these beliefs are therefore lucky. Finally, Lackey (2008, pp. 258-259) has argued that an event can be lucky for a subject even if the subject has direct control over the event.

On the modal account of luck, developed by Duncan Pritchard, an event is lucky if it obtains in the actual world but fails to obtain in most nearby possible worlds where the relevant initial conditions for the event remain the same (2005, p. 128). The modal account treats a fair lottery win as a lucky event since it occurs in the actual world, but fails to occur in most nearby possible worlds where the relevant initial conditions for the event stay the same (namely the lottery is fair). After all, the number of possible worlds where one’s lottery ticket is a loser far outnumber the possible worlds where the ticket is a winner. The modal account also respects
the intuition that it is not a matter of luck that one won the lottery if one has rigged the game. Presumably, if one has rigged the lottery one will win, not just in the actual world, but also in most nearby possible worlds where one remains in control of which ticket wins. Unlike the lack of control account, however, the modal account does accommodate the intuition that it is not at all a matter of luck that the sun rose this morning, since the sun continues to rise in all nearby possible worlds.

The modal account of luck is consistent with empirical evidence regarding how people conceive lucky events. Psychologist Karl Teigen (1995, 1996, 1997, 1998a, 1998b, 2003) found out that success was perceived as more lucky if the possible failure was physically close rather than faraway (for example, when a wheel of fortune stopped in the winning sector which was physically close to a losing sector). Interestingly, Teigen found out that the perceived luck could not be understood solely in terms of the probabilities involved, since subjects would perceive that a ball which landed in a losing sector which was physically near a winning sector involved more bad luck than if the ball landed on a losing sector which was not near a winning sector, all the while recognizing that the chance of a ball landing in a losing sector was constant on each spin of the wheel. Since physical closeness on a roulette wheel is basically modal closeness, Teigen’s results support the modal account of luck.

But even though the modal account of luck offered by Pritchard is able to accommodate our intuitions regarding different cases and is supported by empirical evidence there are dissenters. Lackey (2008, p. 261) has provided a counterexample where a subject is planting a rosebush in location X (and could not have planted it anywhere else) and finds a buried treasure which could not
have been buried anywhere else than in location X. Intuitively, the subject is lucky to have found the buried treasure. After all, she did not know anything about the treasure. However, given how the case is constructed, the subject will find the treasure in all nearby possible worlds where she goes to plant the rosebush, and therefore her finding the treasure is not a lucky event according to the modal account of luck.17

In Essay 1 I argue that the modal account of luck as stated by Pritchard rules certain non-lucky events as lucky. I argue that we must amend the modal account so that we do not restrict our attention solely to the modal profile of the event that occurred in the actual world when evaluating whether the event occurred as a matter of luck. Rather, we need to look at the modal profile of other events as well that are similar in nature and equal in significance to the subject. The reformulated version of the modal account of luck supports a novel reformulation of the safety condition. Even though the modal account of luck does suffer from some problems I do still consider that it is approximately correct. The modal account of luck is largely assumed in essays 1-3, although the results do not depend on the account being true. Rather, the modal account of luck motivates some of the claims that are made in the essays.

Now that we have the modal account of luck before us it should be easy to see why SAFETY is a more promising anti-luck condition that SENSITIVITY. After all, on the modal account of luck a belief is true as a matter of luck just in case it is true in the actual world and false in most nearby possible worlds. A belief is safe just in case it is true in nearby possible worlds where it is believed. Hence a

17 Carter (2010, pp. 530-531) and Pritchard (2015, p. 159) defend the modal account against Lackey’s argument.
belief that is safe cannot be true as a matter of luck if the modal account of luck is the correct account. Can a belief that is sensitive be true as a matter of luck? Yes, it can. Recall that a subject’s belief that \( p \) is sensitive just in case in the nearest possible worlds where \( p \) is false the subject does not believe that \( p \). Since it is possible that a subject will refrain from believing that \( p \) in the scope of the very closest possible worlds where \( p \) is false and believe it falsely in possible worlds that are a bit further away, but still nearby, a sensitive belief can be true in the actual world but false in most nearby possible worlds. Therefore, while both SENSITIVITY and SAFETY have been designed to rule out the kind of luck that is incompatible with knowledge, only SAFETY is backed by a theory of what luck is.\(^{18}\)

Nevertheless, even SAFETY is under threat. Like SENSITIVITY, SAFETY is trivially satisfied if the subject believes in a necessarily true proposition. The reason for this is that a necessary truth is true in all possible worlds and hence it is true in all possible worlds where it is believed. However, as noted above, not all beliefs in necessary truths amount to knowledge. In Essay 1 I argue that in order to deal with cases featuring necessary truths the safety condition must be globalized to a set of propositions in which the subject could easily have believed.\(^{19}\) To be truly safe from error it is not enough that the belief that the subject formed in the actual world could not easily have been false. Rather, the

\(^{18}\) Several virtue epistemologists have argued that the satisfaction of SAFETY (or a condition in its vicinity) is entailed by the satisfaction of a virtue-theoretic condition (Carter, 2016; Gaultier, 2014; Greco, 2010, 2016; Littlejohn, 2014; Sosa, 2015). If this were true then, given the independent plausibility of virtue epistemological theories of knowledge, we would have yet another reason to prefer SAFETY over SENSITIVITY. Sadly this advantage seems to be illusory. I have argued (forthcoming) that the arguments that virtue epistemologists have laid down for the idea that virtue entails safety do not entail the kind of safety condition that we have in mind, when we require a belief to be safe from error in order to be knowledge.

\(^{19}\) Globalized versions of the safety condition have been put forward by Williamson (2009b) and Pritchard (2012a). In Essay 1 I compare my formulation with Williamson’s and Pritchard’s formulations.
subject must be safe as an epistemic agent and this means that she must be safe from forming false beliefs in her inquiry. This point about safety can be made outside the epistemic domain as well. The fact that I am safe from being shot by Smith does not mean that I am safe from being shot period. In order to be truly safe from being shot I must be safe, not just with respect to Smith, but with respect to other potential shooters as well. My proposal is that in order to be properly safe one must be safe from error in one’s inquiry. In order to know it has to be the case that one could not easily have ended up believing in a possible false answer to a question that the proposition one believes in the actual world is an answer to.

Besides being able to deal with cases featuring necessary truths, the globalized version of the safety principle can be applied to the problem of peer disagreement in ways in which the original safety principle could not. The main question in the epistemology of disagreement is what kind of epistemic power disagreement has. This question is usually framed in terms of reasonableness and rationality. Hence the central questions have been can there be reasonable disagreements between epistemic peers, and what ought one do in the face of peer disagreement. In Essay 2 I approach the epistemology of disagreement from a novel angle by considering what kind of epistemic power disagreement has over knowledge, assuming that the safety condition is a necessary condition for knowledge and that knowledge is the norm of belief. The assumption that safety is a necessary condition for knowledge motivates a modal account of epistemic peerhood.

On the modal account of epistemic peerhood, two subjects are epistemic peers regarding a subject matter only if they have roughly the same number of true and false beliefs regarding propositions that belong to that subject matter across
nearby possible worlds in a similar distribution. There are currently two other accounts of epistemic peerhood in the literature. On the evidential and cognitive equality account two subjects are epistemic peers relative to the question whether \( p \) if and only if they have the same body of evidence that is relevant to the question whether \( p \) and they are equally competent and reliable in assessing that evidence.\(^{20}\) A competing account has been proposed by Elga (2007, p. 481) according to which two subjects are epistemic peers on the question whether \( p \) if and only if, conditional on their disagreement they are both equally likely to be mistaken. I have no argument against these accounts of epistemic peerhood. Rather, I merely think that the modal account of epistemic peerhood is much easier to operate on if the safety condition is a necessary condition for knowledge.

Interestingly, it turns out that two subjects who disagree, and really are epistemic peers, cannot satisfy the globalized safety condition if the modal account of epistemic peerhood is correct. After all, given that the subjects acquire roughly the same number of true and false beliefs in nearby possible worlds, and one of them believes falsely in the actual world, it applies to both disagreeing parties that they could easily have believed something that is false. Therefore, even if you are in the right and I am in the wrong neither of us has knowledge, since you could easily have believed what I believe.

Crucially, however, the modal approach to disagreement yields different results in cases of apparent peer disagreement. In such cases the disagreeing parties merely take each other to be epistemic peers, whereas in reality they are not. In such cases the subject who was in the right might satisfy the globalized safety condition and

thus might have knowledge. In Essay 2 I argue that the modal approach to
disagreement offers a fruitful middle-ground between the conformists and non-
conformists, and is actually supported by both conformist and nonconformist
intuitions.

The conformism versus nonconformism debate is one of the central disputes
within the epistemology of disagreement. The conformists claim that we should
give equal weight to our own beliefs and to those held by our peers. They think
that in cases of peer disagreement significant doxastic revision is called for.\textsuperscript{21}
Nonconformism is much more popular nowadays. According to nonconformism,
the mere fact that you disagree with your epistemic peer does not mandate any
revision on either side of the disagreement.\textsuperscript{22} In Essay 2 I argue that conformists
have focused on cases of real peer disagreement and that their view delivers the
correct verdict regarding such cases, while nonconformists have focused on cases
of apparent peer disagreement and that their view delivers the correct verdict
regarding such cases. The modal approach to disagreement explains what is
correct in conformism and nonconformism and what is wrong in both positions.

However, even though the globalized version of the safety condition has several
advantages over its simpler versions, it inherits some severe problems from its
predecessors. Perhaps the most pressing problem for the safety condition is the so-
called generality problem. In order to get to grips with the generality problem we

\textsuperscript{21} Christensen (2007), Elga (2007), Feldman (2007) and Littlejohn (2013) have defended the
conformist position. Lasonen-Aarnio (2013) has argued that no version of conformism is correct
since they ignore what she calls evidential attenuation and evidential amplification. Evidential
attenuation can shield the subject from defeating evidence in such a way that she should not give
equal weight to the opinion of a peer, while evidential amplification makes a subject more
susceptible to defeaters and thus forces her to give greater weight to the opinion of her peer. In
neither case should the disagreeing parties give equal weight to the opinion of their peers (2013, p.
783).

\textsuperscript{22} Nonconformism has been defended by Kelly (2005, 2010), Wedgwood (2010) and Sosa (2010)
and among others.
need to realize that SAFETY as currently formulated is susceptible to certain
counterexamples that can be sidestepped if SAFETY is relativized to the method
of belief formation that the subject uses in the actual world. This weakness is
shared by SENSITIVITY. Nozick (1981, p. 191) himself already noted that modal
conditions for knowledge need to be relativized to the method of belief formation.
For example, suppose that a grandmother sees her grandchild who is in good
health and forms the justified true belief that the child is healthy. But, if the child
had been sick (which could easily have been the case given a recent outbreak of a
virus in the kindergarten) the child’s parents would not have allowed the
grandmother to visit and would have lied to her that the child is in good health in
order not to upset her. Intuitively, the grandmother knows that the child is in good
health when she sees the child, even though the child’s parents would have
deceived her if the child had been sick. Given how the case is set up, the
grandmother’s belief does not satisfy SAFETY or SENSITIVITY. After all, there
is a nearby possible world where the grandmother believes that the child is in
good health while she is sick and in the nearest possible worlds where the child is
sick the grandmother will continue to believe that the child is in good health.

Crucially, if SAFETY and SENSITIVITY are relativized to the method of belief
formation that the subject uses in the actual world, the grandmother will satisfy
both conditions. There are no nearby possible worlds where the grandmother ends
up believing falsely that the child is in good health by seeing the child. Therefore,
the method of belief formation that the subject uses in the actual world must be
kept constant in all relevant possible worlds when evaluating whether the belief is
safe or sensitive.
But the method relativization of the conditions raises a new problem: how should we determine the method of belief formation that the subject used in the actual world to form her belief? Given that all (non-innate) beliefs are tokens of a certain cognitive process or method, and that each token can be labelled under multiple different types of methods, it is paramount to explicate how the token processes are to be typed. The extension of safe and sensitive beliefs will vary greatly depending on how the tokens are typed. This is essentially the generality problem for process reliabilism.23

Following Conee and Feldman (1998, p. 3), epistemologists have thought that in order to solve the generality problem the reliabilist has to provide a formula that picks out a unique process type in each case. Reliabilists have tried to meet this demand by trying to offer necessary and sufficient conditions for when two token processes can be treated as belonging under the same type.24 I believe that this strategy of answering the generality problem is misguided. In Essay 1 I offer a solution to the generality problem that becomes available once we recognize that if our aim is to provide an account of what knowledge is, it is enough to be able to tell when two knowledge-conducive processes belong under the same type. I argue that it is much easier to answer this question.

In order to avoid obvious circularity in the analysis of knowledge I offer an account of knowledge-conducive methods of belief formation that is motivated by virtue epistemological considerations. I argue that only virtuous methods of belief formation can produce knowledge and that therefore it is harmless to restrict our

---

23 The fact that reliabilist analyses of knowledge are susceptible to the generality problem was first noted by Goldman (1979) and has later been pursued notably by Conee and Feldman (1998).
attention to virtuously formed beliefs when evaluating whether a given belief satisfies the safety condition. Epistemic virtues are stable dispositions seated in the agent to attain true beliefs and avoid false regarding a certain field of propositions while in certain conditions (Sosa, 1991, p. 284). Very roughly, in the case of the virtue of good eyesight the physical basis of the disposition would consist in the rods and cones and the visual cortex, the field of propositions would consist of propositions in which the virtue could produce beliefs in ([that is red], [that is round]) and the conditions would be adequate lighting conditions, medium distance etc. In Essay 1 show that by relativizing the safety condition to virtuous methods of belief formation that the subject uses in the actual world the safety theorist is able to provide a promising solution to the generality problem. 

This version of the safety condition motivates a ‘minimal’ virtue epistemological theory of knowledge that is committed to the idea that epistemic virtues are necessary for knowledge. It rejects, however, the idea that in all cases of knowledge, one’s cognitive success, i.e. the acquisition of a true belief, is attributable to one’s cognitive virtues. In Essay 3 I argue against virtue epistemological theories of knowledge that are committed to the latter of these

25 While the idea that knowledge is always gained through virtuous methods of belief formation is widely accepted Lasonen-Aarnio (2010) has argued that knowledge can be gained even through bad epistemic strategies. However, even on her view the method via which the belief was gained has to be reliable in the particular circumstances in which it was used if the belief is to constitute knowledge. While the view advocated by Lasonen-Aarnio is in tension with the view presented here, I believe there is a way to reconcile them, since the fact that one formed a belief through a virtuous method of belief formation does not obviously entail that one must be “managing one’s beliefs through the adaption of policies that are generally knowledge conducive” (Lasonen-Aarnio, 2010, p. 2). Exercising a virtuous method of belief formation need not entail that one is thereby disposed as an epistemic agent to manage one’s belief-forming processes in a way that is knowledge-conducive. While the position advocated by Lasonen-Aarnio might be compatible with the one pursued here, I would rather go in the opposite direction and argue that knowledge cannot be acquired by subjects who are not epistemically virtuous. This move would allow the proponent of the safety condition to accommodate our intuitions in defeat cases, without resorting to an error theory, like Lasonen-Aarnio does.

ideas. I argue that several high-profile virtue epistemologists are committed to the idea that one’s cognitive success is attributable to one’s cognitive virtues just in case the fact that one formed one’s belief through the exercise of one’s cognitive virtues enhances the safety of one’s belief. Call this the contribution thesis. I demonstrate that if the contribution thesis was true, virtue epistemologists would be able to deal with some of the most pressing arguments against virtue epistemological theories of knowledge. Namely, they would be able deal with both cases featuring environmental epistemic luck and with cases of testimonial knowledge and accommodate the epistemic dependency thesis. However, I also show that virtue epistemological theories of knowledge that are committed to the contribution thesis are ultimately untenable, since there are clear cases of knowledge where one’s cognitive virtues do not enhance the safety of one’s belief.

While the generality problem and the problem concerning necessary truths do not undermine the idea that safety is necessary for knowledge, several authors have provided arguments that seek to establish that conclusion. Even though I do not argue for the claim that a safety condition is necessary for knowledge in the essays that this dissertation is comprised of, the conclusions that one would be able to draw from this dissertation would be more interesting if safety was necessary for knowledge. Since I do not engage with the arguments that purport to

27 Lackey (2007, 2009) has argued that virtue epistemological accounts of knowledge cannot accommodate cases of testimonial knowledge while delivering the verdict of ignorance in cases featuring environmental luck. Pritchard (2010, 2012a) has argued for the same conclusion with respect to robust virtue epistemological theories of knowledge that contain only virtue-theoretic conditions, but he maintains that weaker virtue epistemological theories can accommodate both cases. Kallestrup and Pritchard (2013) have argued that robust virtue epistemological theories of knowledge cannot accommodate the idea that knowledge can be dependent on factors that are completely external to one’s cognitive agency and therefore is unable to deliver correct verdicts regarding certain cases.
show that safety is not necessary for knowledge in the essays, it is proper to take a
stance on them here.

The arguments that have been presented against the necessity of the safety
condition rely on the counterexample method. Neta and Rohrbaugh (2004),
Comesaña (2005), Baumann (2008) Kelp (2009) and Bogardus (2014) have
provided thought experiments where the subject’s belief does not seem to satisfy
the safety condition and yet the subject’s belief seems to be knowledge. ‘Seems’
is the right choice of word here since many epistemologists have challenged the
idea that the thought experiments feature either a safe belief or knowledge. For
example, Bogardus (2014) has argued that in the putative counterexamples put
forth by Neta and Rohrbaugh, Comesaña, Bauman and Kelp, the subject actually
satisfies the safety condition and therefore they are not counterexamples to the
necessity of the safety condition. Broncano-Berrocal (2014) and Pritchard (2016,
p. 9) have argued that the counterexample that Bogardus presents suffers from the
same defect. Williamson (2009a) has also argued that the safety condition is
satisfied in the case offered by Neta and Rohrbaugh. Goldman (2009, pp. 79-80)
notes that the case presented by Neta and Rohrbaugh is analogous to the famous
barn façade case and therefore seems to be a case of ignorance rather than a case
of knowledge. Pritchard (2012b, pp. 181-183) argues that the cases presented by
Neta and Rohrbaugh and Kelp are not cases of knowledge and provides an
explanation why they might seem to be cases of knowledge.

In short, to say that there is no consensus regarding whether there are genuine
counterexamples against the necessity of the safety condition is a gross
understatement. Rather than going through immensely complicated intuition
mongering, I will here just state some structural features that are shared by the
cases offered by Neta and Rohrbaugh, Comsaña, Bauman and Bogardus and refer to Luper who has argued that such cases are best conceived as borderline case of knowledge. For brevity’s sake I will here present only the case offered by Neta and Rohrbaugh, which goes as follows: S is drinking a glass of water which she has just poured from the bottle. Standing next to S is a happy person who has just won the lottery. Had this person lost the lottery, she would have polluted the water with a colourless, odorless, tasteless toxin, but since she won the lottery she does nothing of the sort (2004, pp. 399-400).

According to Neta and Rohrbaugh, S knows that she is drinking water in the case, even though her belief is unsafe. After all, there are many nearby possible worlds where she believes that she is drinking water while her belief is false, since her ‘friend’ has poisoned the water.

Luper (2006) argues that the above case features ‘restorative rigging’ and that it therefore differs in important respects from clear-cut cases of ignorance that feature only ‘deleterious rigging’. Classic Gettier cases, such as the clock case, feature deleterious rigging. In such cases, the subject is using a method of belief formation that is normally reliable, but which happens to be unreliable in the peculiar circumstances in which the subject finds herself in. In essence, the reliability of the method is undermined by introducing further details to the case, in the clock case by telling that the clock has stopped. Cases like the one offered by Neta and Rohrbaugh differ from Gettier cases in that after introducing details that rig the case deleteriously the reliability of the method of belief formation is restored by some further added details (in the above case by telling that the person won the lottery). Figuratively speaking, the deleterious rigging ‘pulls’ the worlds in which S is deceived closer than they would have been if there had not been a
person with malicious intentions next to S, while the restorative rigging ‘pushes’ them back out. Depending on how the details of the case are spelled out the deceit-worlds might be pushed far enough for the safety condition to be satisfied (Luper, 2006, p. 164). Luper argues that cases that feature restorative rigging are best conceived as borderline cases of knowledge. This conclusion is further strengthened by the fact that several epistemologists have suggested that cases featuring restorative rigging are in fact cases of ignorance, and drawn the conclusion that the safety condition is not sufficient for knowledge.\(^{28}\) Given that epistemologists have diverging intuitions regarding these structurally similar cases, the fact that the safety condition rules such cases as borderline cases of knowledge should be seen as a virtue of the safety condition. The inherent vagueness of the safety condition is unproblematic if it corresponds to the vagueness we find in our knowledge attributions, and this seems to be the case (Williamson, 2000, p. 100).

This does not mean that we should blindly accept the safety condition as a necessary or sufficient condition for knowledge. Hand-waving arguments for one side or the other should not be used to determine whether safety is necessary for knowledge. On the contrary, there are genuine doubts as to whether the satisfaction of the safety condition is necessary for knowledge possession. But even though we might rightfully doubt whether safety is necessary for knowledge, we should recognize that having epistemically safe beliefs is of great value. We need to minimize the risk of false belief if we are to succeed in our practical endeavours. This thesis argues only indirectly for the necessity of the safety condition. It does so by highlighting how the concept of epistemic safety can be

---

put to use in different epistemological debates with interesting and plausible results. One might say that I have put the carriage before the horse, but sometimes you should do this, if only to see what comes of it. All in all, I consider that there are better reasons to think that safety is necessary for knowledge than to think that it is not. However, if it turned out that safety is not necessary for knowledge this thesis would not be undermined. Epistemic safety is in itself a philosophically and practically important phenomenon, and unearthing what it is, how it is related to epistemic virtues and whether we can believe safely in the face of peer disagreements are philosophical discoveries of value.
References


