

Abstract

Action, Knowledge, and Authority: Variations on a Reidean Theme

Juan S. Piñeros Sánchez

2019

The central aim of this study is to develop an account of practical knowledge, knowledge acquired on the basis of how we affect the world, rather than how the world affects us. I approach the question from within the Reidean tradition in epistemology. A central tenet of this tradition is that there are a set of privileged sources of information that give us knowledge of the world in a direct (i.e. non-inferential) way. The traditional list includes only theoretical sources, but I argue that it should also include two practical ones: the will (the faculty in virtue of which we intend and act intentionally) and authority. To defend this thesis, I develop a virtue-theoretic account of epistemic warrant, and I use this account to show how testimony can give us knowledge in a direct yet personal way.

Chapter 1 focuses on a widely endorsed Anscombean thesis about agential knowledge (the knowledge agents characteristically have of their own actions). According to this thesis, when an agent acts intentionally, she necessarily knows what she is doing. I defend a version of Timothy Williamson's anti-luminosity argument to show that this thesis is untenable for limited creatures like us. Indeed, the objection shows that no such necessary connection could obtain between knowledge and any practical entities (such as intentions).

Chapter 2 argues against cognitivist views that draw a necessary connection by appeal to weaker cognitive entities (such as beliefs) rather than knowledge. I argue that these

views suffer from counterexamples, and that we can provide a principled recipe to generate such counterexamples. We should thus reject cognitivism.

Chapter 3 considers the most natural alternative to cognitivism: inferentialism. On an inferentialist view, we gain knowledge of our actions on the basis of an inference based on knowledge of our intentions. I argue that inferentialism cannot account for the characteristic directness and spontaneity of agential knowledge. I also identify four desiderata that an account of agential knowledge must meet based on considerations from the first three chapters.

Chapter 4 lays the epistemic groundwork for the remaining chapters by developing an account of direct warrant, the kind of warrant possessed by beliefs that are not based on other beliefs. I follow virtue epistemologists in holding that it is sufficient for a belief to be warranted if it is competently held, and I develop an account of competence in terms of guidance. I illustrate the advantages of the account for the case of perceptual knowledge.

Chapter 5 presents my positive account of agential knowledge. It has two central components. First, I argue that the beliefs that constitute such knowledge are intentions. However, unlike traditional cognitivists, I hold that the identity between beliefs and intentions holds only at the level of token mental states: intentions and beliefs are different *kinds* of states, but a single state can be of both kinds. Second, by appeal to the account from Chapter 4, I argue that beliefs about our actions (like, the belief that I am writing) are self-warranted, like the belief that I am thinking. I show how the resulting account of agential knowledge meets the four desiderata identified in Chapter 3.

Chapter 6 argues that practical knowledge encompasses more than just agential knowledge. There is a second species of practical knowledge that we have in virtue of

exercising practical authority, which I call ‘authoritative knowledge’. The chapter presents two arguments to the effect that this knowledge is direct and practical.

Chapter 7 closes by defending an interpersonal account of testimonial knowledge. I argue that to be able to acquire knowledge from a speaker directly while regarding her as a person, the listener must engage in a second-personal transaction of the sort that we engage in when we make a promise. A central advantage of the view, over other interpersonal accounts in the literature, is that it gives a specifically epistemic role to the second-personal reasons that come into play when testimonial transactions take place.

Action, Knowledge, and Authority: Variations on a Reidean Theme

A Dissertation
Presented to the Faculty of the Graduate School
of
Yale University
in Candidacy for the Degree of
Doctor of Philosophy

by
Juan S. Piñeros Sánchez

Dissertation Director: Gideon Yaffe

May 2019

© 2019 by Juan S. Piñeros Sánchez

All rights reserved

Table of Contents

Acknowledgments	iii
Introduction	1
Chapter 1: Practical Knowledge and Luminosity	8
1. (PKP) and Presumed Counterexamples	10
2. Practical Anti-Luminosity	19
3. Reconsidering the Non-Accidental Condition	34
Chapter 2: Against (Traditional) Cognitivism	39
1. Arguments for Cognitivism	42
2. Counterexamples Generalized	44
3. Beyond Standard Cognitivism	49
Chapter 3: Against Inferentialism	52
1. Inferentialism	53
2. Chatty Charlie's Assertions	57
3. Alienated Knowledge	63
4. Taking Stock	72
Chapter 4: Knowledge through Virtuous Performances	75
1. Foundationalism and the Need for a Better Reliabilism	77
2. Warrant by Competence	80

3. Conclusion	88
Chapter 5: The Will as Epistemic Faculty	90
1. A Moderate (Token-Level) Cognitivism	91
2. Self-Warrant Through Competence	105
3. Conclusion	115
Chapter 6: Authoritative Knowledge	117
1. Authoritative Knowledge Further Explained	119
2. Direct and Indirect Models	122
3. Arguments for the Direct Model	125
4. Direct Knowledge in Three Virtuous Steps	132
5. Implications and Complications	136
Chapter 7: Freedom and Assertion	139
1. Interpersonal Views and Their Place in the Literature	141
2. Lackey's Dilemma	149
3. Testimonial Knowledge and the Aim of Assertion	153
4. The Epistemic Function of Assurance	158
5. Evidentialism and Epistemic Partiality	166
References	171

Acknowledgments

I owe thanks to many people for their help throughout the intellectual journey that culminated in this work. I would probably not be doing philosophy today if it weren't for Leo Ehlers, who introduced me to the subject in high school, and kindled my interest in Aristotle. In addition to Ehlers, I want to thank especially Ms. Elliott, and Mr. Rogers for their patience and encouragement as I mastered a new language.

When I moved on to the University of Toronto to pursue my BA, Ehlers told me to 'find good people, and let them educate you'. I found many good people at U of T who were willing to educate me. They include Don Ainslie, Joe Goering, Doug Hutchinson, Mark Kingwell, Elmar Kremer, and Giulio Silano. The list also includes the late Joe Boyle, who introduced me to Anscombe's philosophy of action. I miss him, and wish I could discuss the ideas in this work with him.

In my fourth year at U of T I also took a course on personal identity with Jennifer Whiting. Jen once told me that a conversation with Rogers Albritton completely changed the shape of her dissertation. I was very lucky to have not just one but many conversations with Jen that radically shaped my philosophical outlook, which no one has influenced as much as her. It is her refrains about philosophical methodology that are at the back of my head whenever I do philosophy. I cannot thank her enough. I am also proud to call her not just a mentor but a friend.

After my BA, I applied for the MA program in philosophy at U of T. I think partly on Jen's advice they offered me a place in their PhD program, with the option of applying elsewhere later on. The years at Toronto were full of intellectual stimulation and I have many teachers and colleagues to thank for conversations and instruction. The teachers

include Rachel Barney, Deborah Black, Phil Clark, Imogen Dickie, Lloyd Gerson, Brad Inwood, Bernie Katz, Martin Pickavé, Sonia Sedivy, Andrew Sepielli, Denis Walsh, and Jessica Wilson. The colleagues include Dominic Alford-Duguid, Mike Arsenault, Juan Pablo Bermúdez, Willie Costello, James Davies, Jason DeRoche, Mark Forney, Dan Ioppolo, Marta Jimenez, Kevin Kuhl, Geordie McComb, Parisa Moosavi, Daniel Rabinoff, Bryan Reece, Luke Roelofs, Ljiljana Stanić, Jacob Stump, Trevor Teitel, Johanna Thoma, Ben Wald, and Steve Woodworth. I would also like to thank Anita Di Giacomo, Mary Frances Ellison, and Margaret Opoku-Pare for all their help. Finally, at Toronto I also met Allison Piñeros Glasscock, Reier Helle, and Robbie Howton. More on those three soon.

For various reasons, I decided after two years to apply to a few other graduate programs. I was actually in New Haven when I received my offer from Yale. I was ecstatic, and accepted soon after, before the official visiting days. Despite the hardships, these have been the happiest years of my life, and my philosophical views have continued to mature. I have learnt so much from so many people here. Michael Della Rocca taught me how to take a philosophical idea and run with it; Keith DeRose introduced me to epistemology and planted Reidean seeds in my head; Verity Harte and Brad Inwood showed me how to read texts with outmost seriousness and care; and Josh Knobe and Zoltan Gendler Szabó taught me how to use empirical and linguistic work to throw light on philosophical questions. Zoltan also provided much needed encouragement and advice for the job market. I have discussed some of the ideas included here with all of these people; I sometimes feel as though at Yale I have a much larger committee that includes them all. Finally, Timothy Williamson began lecturing at Yale every spring semester in my third year. Apart from Anscombe's works, Tim's writings in epistemology are probably the biggest intellectual

influence on the ideas presented below. I have gained immensely from the discussions about many of them with him.

When I started in the program, Vicki D’Agostino and Pat Slatter helped me get settled in the department. When they left, I have relied for so many things—like everyone else here—on Julie Niemeyer, Monique Boney, and Jasmine Williams. Thank you!

I have been asked several times about my favourite part of doing philosophy. Without a doubt, the answer is: the conversations. I love discussing philosophical ideas, and at Yale I have found many friends with whom to do so. Many of these conversations took place in 98 Avon, a place that was eventually shared between Matt Leisinger, Reier Helle, and I. Matt and Reier were often the first to hear my ideas in the course of ‘boardgame night’. I will always look fondly on the memories from these years. Many of these memories include Emily Kress and (eventually) Eric Guindon. At some point, my dominance at boardgames made it a necessity to find someone else who could provide a real challenge. This, I believe, is how we became friends with Jo Demaree-Cotton and Maxime Lepoutre, who joined the boardgame crew. There were also occasional guest appearances by Harry Alanen, Julia Borcharding, Robbie Howton, and Lea Schroeder. Beyond the philosophical exchanges, these friendships have added infinite amounts of joy to my life. It is a testament to our friendship that I was comfortable asking Reier, Jo, and Maxime (as well as my family) to proofread parts of this work on very short notice. They saved me from many mistakes. Those that remain are of course their fault.

Some of the ideas that appear here developed in the course of reading groups that I ran with Mike Deigan (on puzzlement and inquiry), Daniel Moerner (on philosophy of action), Jessie Munton (on perception), Yuan Yuan (on authority), as well as MDR and

Dan Greco (on Sosa's virtue epistemology). I was very lucky that among my friends were such outstanding interlocutors to work on topics I was interested in. In addition to the aforementioned colleagues, there are several others who have provided substantial intellectual and emotional support throughout these years including Mario Attie, Joanna Blake-Turner, Chris Blake-Turner, Mike Burton, Tim Clarke, Jennifer Daigle, Brian Earp, Bridger Ehli, Sam Elgin, Dan Ferguson, Hugo Havranek, Arthur Lau, Matt Lindauer, Mark Maxwell, Moya Mapps, Stephen Ogden, Elisabetta Pellegrino, Armando Perez-Gea, Daniel Putnam, Will Ratoff, Ajay Ravichandran, Evan Rodriguez, Jake Rohde, Paul Schilling, Kate Stanton, Jonathan Vertanen, Dylan Vollans, Marissa Wallin, and Alex Zhang. Finally, my experience in New Haven has been enriched by my friendships with many non-philosophers, including Ignacio Quintero, Kaitlyn and Chris Bartley, Esther Florsheim, Humberto González Chávez, Rachel Love, Juan Penagos Zuluaga, Steph Rohde, Oscar Soler Sanchez, Nate Upham, Jennifer Weintritt and Evan Johnson.

In the course of writing this thesis, I received incisive written feedback from several people outside my committee. I thank Sarah Buss, Michael Della Rocca, Ben Lennertz, and Antonia Peacocke for comments on Chapter 1; John Schwenkler for comments on Chapter 2; Sarah Buss, Sarah Paul, and John Schwenkler for comments on Chapter 3; Sarah Paul for comments on Chapter 5; Jewelle L. Bickell, Christopher M. Cloos, Zoltán Gendler Szabó, Daniel Moerner, Jessie Munton, and Yuan Yuan, for comments on Chapter 6; and Mark Boespflug, Sarah Buss, Tom Cook, Teresa Bruno-Niño, and Keshav Singh for comments on Chapter 7.

The previous paragraph is evidence of the support that my work has received from two people in particular. The first is John Schwenkler, with whom I share a broad outlook

in philosophy of action, and likely for that reason has raised some of the sharpest objections to my views. The second one is Sarah Buss, who I met at a conference in 2016. At that time, I had no idea how one was even supposed to approach someone whose work one greatly admired. Sarah made it easy by demanding that we talk over lunch about the paper I had presented. She raised an objection to my paper that eventually led to a complete revamping of the project that is now Chapter 7. This was for the best. Sarah has an enviable ability to get at what's deep and important in philosophy. She has provided invaluable feedback on my work, as well as tremendous encouragement in the past couple of years.

Now on to my committee.

At the time when my central focus was on ancient philosophy, I had this dream of going to Oxford to work with David Charles. I couldn't believe my luck when he accepted an offer from Yale the year I started here. David's influence on my views on action theory permeates this dissertation. I also look to him as an exemplar of how to conduct oneself in philosophical discussions, always kindly and truthfully (with ἀλήθεια). The process of learning from him is ongoing (even as we teach together).

The two people who led my proseminar ended up in my committee. The first is Dan Greco, who iterates from excellent question to excellent question every time I discuss my ideas with him. He has been an invaluable resource on contemporary epistemology, and he has kept me safe from several mistakes. Like many other graduate students in the program, I have also benefited immensely from his advice on grad school and professionalization.

The second person leading the proseminar was Steve Darwall. Like many others, I regard his work on interpersonal normativity as one of the most important contributions to moral philosophy of the past century. The influence of his work is particularly evident in

the last chapter, but it goes well beyond. Steve is also a role model of work ethic. I do not know how he manages to live up to the highest expectations one might have of a committee member given all his other commitments. I am convinced that he owns a time-turner.

One reason it was easy to accept Yale's offer before visiting was that I knew Jason Stanley was moving here. At that point, I had read only bits of *Know How*, but, despite the high stakes, that was enough to know I wanted to work with him. I have learnt from Jason to think big and what that takes to do so, while maintaining the highest standards of rigour and clarity. His enthusiasm for my work has also made it much easier to pursue this project.

Every major idea in the dissertation has been sharpened and improved through my discussions with Gideon Yaffe. I would look forward to each of our meetings eagerly, wondering how many of my ideas would survive his criticisms. Those that made it usually started out in a messy form that would take proper shape in the course of our discussions. More than anyone I know, in both his writings and conversation, Gideon can get at the heart of a messy idea and state it in clear terms from which it can be properly assessed. Gideon manages to be an extremely kind interlocutor, without ever lowering the bar for his students. I cannot adequately express how amazing a supervisor he has been. Perhaps the best I can say is that as I move on to the stage of my career, I will always attempt to be as good a mentor for others as he was for me (however futile this might be).

Throughout my life since moving to Canada I have been supported by some people who once in a while are willing to discuss philosophy with me. Alex Belanger, Christian Diamond, Miles Krauter, Paul Fiorentino, Karla Wobito, Billy Davidson, Russell Jackson-Wyatt, Andrea Zuccolin, Taylor Farley, Ryan Peardon, Nicole lemon, Ian Lemon, Ashleigh

Altomare, Elizabeth-Marianne Violante, Nicolás Barake, Simón Ucrós, José Salomón Marun Helo, and Andres Felipe García: thank you for your friendship.

The Glasscocks have made me feel like part of their family well before I officially became part of it. Thanks for all your love Shannon, Lori, Bethany, Cassio, Luciano, Caio, Cameron, Angela, Daniel, Jessica, Jordan, Elaina, Jeremy, Zipporah, Faerynn, Cody, Gavin, (and Chesterton!).

¿Qué decir de mis hermanas, mis compañeras desde niños, y mis amigas para siempre? María José y Luisa: Gracias por su alegría, su constante apoyo, y la felicidad de ser su hermano. Nunca es más que un hasta luego, nunca es más que un nuevo adiós. Papás: Creo que nunca se imaginaron que terminarían con un hijo filósofo. Esto es prueba de la libertad que siempre nos han dado para encontrar nuestro propio rumbo, apoyándonos durante todo el camino. Se que no lo digo mucho, pero también sé que lo saben: Los quiero mucho, familia.

In the autumn of 2010, when I was starting my postgraduate work, I struck a conversation with someone who was starting her MA at Toronto. We talked about Borges, Boethius, and the word 'slope'. The conversation, more or less on these topics, continues to this day. Allison: I love you with all my heart. I hereby dedicate this work to you.

To Allison, who once went on a date without knowing it.

Introduction

The notion of practical knowledge is explored in the writings of medieval writers like Avicenna and Aquinas, and may go back to Aristotle. However, it did not receive much attention after the medieval period until it became one of the central themes in G.E.M. Anscombe's *Intention* (1958), a work that Donald Davidson called, "the most important treatment of action since Aristotle". I agree with his assessment, though I would be tempted to delete the 'since' clause.

To understand the importance of Anscombe's work on practical knowledge, we need to set it against the frame of the Cartesian philosophy of mind that was still dominant at the time when Anscombe wrote her monograph. Descartes, as is well known, posited a sharp metaphysical division between the mental and the worldly. He argued that part of the difference lies in the way in which the mind and the world are known to us: we have a distinctively direct and authoritative knowledge of our minds that contrasts with our indirect and fallible knowledge of the world. In *Intention*, Anscombe challenges this picture by arguing that we have a similarly direct and authoritative knowledge of our intentional actions, paradigm instances of worldly occurrences. However, if worldly occurrences bear this mark, the Cartesian distinction between mind and body collapses.

Anscombe clearly saw the importance of this point for a theory of mind, but her main interest in the topic was probably ethical. She had seen Oxford award a degree of recognition to President Truman for his actions in WWII. She (along with a few other members of Oxford that included Philippa Foot) protested vehemently: they saw that, as the man responsible for the nuclear attacks in Nagasaki and Hiroshima, Truman was responsible for genocide. Anscombe believed that part (but only part) of the reason why

people failed to recognize this action for what it was, was the faulty moral psychology that naturally arises from Cartesian dualism. From a Cartesian standpoint, we must distinguish between the intention, understood as a purely mental occurrence, and the action, understood as mere effects of this intention. Given such an outlook, it is natural to hold (i) that insofar as the agent is responsible for something, she is mainly responsible for her intentions, and (ii) that insofar as she is responsible for what we would call her ‘actions’, she is no more responsible for them than for any other effects of her intention that she could have foreseen (after all, they relate to the intention in fundamentally the same way). Since Truman intended to end the war by dropping the bombs, and since it was taken that ending the war had prevented many more deaths than were caused by the nuclear bombings in Japan, it follows from (i) and (ii) that Truman’s recognition was well-deserved. For he was, quite literally, a life savior.

Anscombe sought to stop this repugnant train of thought on its tracks by calling into question the account of the connection between intention and action that it presupposes. Centrally, she contends that we do not know about our actions in the same way as we know about other consequences of what we do or intend; indeed, that we know about them in a fundamentally different way than how we know about most other occurrences in the world. We have practical knowledge of our actions, a knowledge that, taking a phrase from Aquinas, Anscombe describes as “the cause of what it understands”. Our relation to our actions is thus fundamentally different from our relation to their side-effects. Truman was fundamentally responsible for the massacres in Japan in a way in which he was not responsible for other side-effects of the bombings.

The final motivation for Anscombe is epistemological. As Anscombe noted in *Intention*, philosophers after the medieval period seemed to have a purely theoretical understanding of knowledge. Indeed, to this day writings in epistemology continue to treat theoretical knowledge as the central topic of study (few introductory epistemology classes or textbooks even discuss the notion of practical knowledge). Yet, practical knowledge is pervasive in our lives: it is how I currently know that I am typing, that I will stand up in five minutes to take a break, and that I will print these pages before the end of the week to submit them to the Registrar's office; and it is how you know that you are reading them (when you do), and that you will stop (whenever you plan to). One might thus wonder how an exclusive focus on theoretical knowledge could not but result in an excessively narrow, and thus distorted epistemological outlook.

It is in terms of these three motivations that we should understand Anscombe's contention that whenever an agent acts intentionally, she has (practical) knowledge of what she is doing. Working backwards, this thesis (1) highlights the pervasiveness of practical knowledge in our lives; (2) it could serve as the basis for defending the view that what is done intentionally and what is merely foreseen should receive different treatment in our moral philosophy; and (3) it challenges the Cartesian picture by including worldly entities (actions) among those to which we have a certain privileged access.

However, it seems to me that Anscombe's challenge to Cartesian philosophy of mind ultimately falls short. While rejecting his thesis that we have privileged access *only* to our own minds, Anscombe seems to share Descartes's conviction that the contents of our minds are always epistemically accessible to us. Her thesis about the connection between intentional action and practical knowledge can thus be seen as the kind of view

that Descartes held about thought in general, but applied to practical thought and action. Yet, recent developments in philosophy and psychology make this Cartesian thesis highly unattractive, as they suggest that we are often wrong about what we think. If this is true of even the most inward thoughts, it stands to reason that it is true of intentional actions. This is the central claim of *Chapter 1*, where I develop one of the most influential attacks on this Cartesian outlook—Timothy Williamson’s (2000) anti-luminosity argument—and show that it provides a particularly strong objection to Anscombe’s thesis, which has undergone something of a revival in recent years.

Much of the rest of the study is an attempt to provide an account that preserves the distinctive features of practical knowledge and its connection to intentional action without presupposing the aforementioned epistemological thesis. It emerges in the course of this study that practical knowledge actually comes in two species, what I call ‘agential’ (knowledge of our own actions) and ‘authoritative’ (knowledge on the basis of authority). And the epistemic framework developed to understand these two varieties of practical knowledge throws light on testimonial knowledge and its interpersonal nature.

A natural way to preserve a connection between agential knowledge and intentional action is to weaken the thesis that having an intention or acting intentionally entails having *knowledge* of our actions, holding instead that it entails having a weaker epistemic state, like belief. This is the central thesis of traditional cognitivism. The problem is that outstanding formulations of this thesis are subject to counterexamples. Moreover, as I argue in *Chapter 2*, there is no principled way to avoid these counterexamples, since they arise from a deep connection between cognitive states and evidence: with few exceptions, it is always possible to have overwhelming evidence to have a different epistemic state

than one in fact has. As such, we can formulate a recipe (à la Zagzebski (1994)) for arriving at such counterexamples, showing that cognitivism in its traditional form is untenable.

In light of these problems, we could be driven to an inferentialist account of agential knowledge. On the most plausible version of this view, this knowledge acquired on the basis of an inference from knowledge of our intentions. However, I argue in *Chapter 3* that inferentialism cannot account for two central features of agential knowledge. First, inferentialism cannot account for the characteristic directness of this knowledge, as is shown by examining the connection between intentions and assertoric expressions of intention. Second, inferentialism conflicts with the characteristic spontaneity of agential knowledge; that is, it conflicts with the fact that this knowledge is available only from the agential standpoint from which our actions are seen as up to us.

On the basis of the considerations from the first three chapters, I close *Chapter 3* by outlining four desiderata for a successful account of agential knowledge: (1) it must explain its non-accidental connection to intentional action; (2) it must explain why we can express our intentions assertorically; (3) it must explain the directness of agential knowledge; and (4) it must explain its spontaneity.

I approach the task of providing an account that meets these desiderata from within a Reidean epistemological framework. Central to this framework is the contention that there is a set of privileged sources of information (like perception and memory) that gives us direct (i.e. non inferential) knowledge of the world. The plausibility of this framework depends on the possibility of providing an account of how such direct knowledge is so much as possible. In *Chapter 4* I defend such an account, developing the virtue-theoretic framework developed by Ernest Sosa, according to which one is warranted in forming (or

maintaining) a belief if one does so competently. I provide a definition of competence in terms of guidance, and use it to explain why perceptual beliefs are directly warranted, and, so, provide us direct knowledge of the world.

Chapter 5 presents my positive account of agential knowledge. It consists of two central theses. First, I argue that the beliefs that constitute such knowledge are intentions. However, unlike traditional cognitivists, I hold that this identity obtains only at the token level: intentions and beliefs are different *kinds* of states, but a single state can be of both kinds. Yet, this view seems to face a dilemma: it entails that agential knowledge is either warranted on the basis of practical reasons or on the basis of no reasons. Either way, it would seem to not be genuine knowledge. To avoid the dilemma, I use the account from *Chapter 4* to defend the view that beliefs about our actions (like, the belief that I am writing) are self-warranted states, like the belief that I am thinking. I show how the resulting account of agential knowledge meets the four desiderata identified in *Chapter 3*.

Discussions of practical knowledge in the literature focus entirely on knowledge of the agential kind, the knowledge agents have of their own actions. However, I argue in *Chapter 6* that there is a second species of practical knowledge that we can acquire on the basis of exercising authority. I call it ‘authoritative knowledge’. Examples include knowing the name of one’s ship just on the basis of baptizing it, or the whereabouts of one’s subordinates just on the basis of ordering them to be somewhere. I argue for this view both by appeal to my account of non-inferential warrant, but also by presenting independent arguments for the claim that authoritative knowledge is non-inferential.

Finally, in *Chapter 7* I apply my account of non-inferential warrant to the case of testimony to defend the view that testimonial knowledge is a second-personal

phenomenon. The argument begins by identifying a tension between two central features of testimonial knowledge: its directness and its personal nature. To be direct, we must assume that the speaker's assertions aim at providing us with knowledge; but regarding the speaker as a person requires us to recognize that she is free to set her own aims for her actions, including assertions. The tension can only be resolved if the speaker freely binds herself to the aim of providing us with knowledge; and to do so, I argue, the speaker and the listener must engage in a second-personal transaction.

Chapter 1: Practical Knowledge and Luminosity

When a person knows something, the world is non-accidentally as she thinks; and when a person acts intentionally, the world is non-accidentally as she intends. Knowledge and action thus yield a distinctive harmony between mind and world. Partly on the basis of these parallels, some philosophers, including Aristotle, have argued that there is an important link between a person's intentional actions and her knowledge of them;¹ that when a person acts intentionally, she must know what she is doing.² This is one of the central theses of Anscombe's seminal *Intention*, where she argues that in cases where an agent lacks knowledge of what she is doing, "what happens does not come under the description—execution of intentions".³

We can put the Aristotle-Anscombe thesis about the connection between action and knowledge more formally as follows:

PRACTICAL KNOWLEDGE PRINCIPLE (PKP): (Necessarily) If an agent is Φ ing (intentionally and under that description), she knows that she is Φ ing (intentionally and under that description).⁴

¹ For elucidative accounts of the parallels in Aristotle, see Moss (2014) and the papers by Allen and Charles in the collection edited by Henry and Nielsen (2015). See also my Piñeros Glasscock (forthcoming) on this point. For a recent extended treatment of the parallels, see Williamson (forthcoming).

² See *NE* 1111a23-4; *EE* 1223a15-7. It might be thought problematic to ascribe this thesis to Aristotle since, as Anscombe (1981 [1965]) noted, it is unclear whether he has our notion of *intentional*. However, in the cited passages Aristotle is listing the cognitive conditions for *voluntary* (ἐκούσιον) action, and it seems safe to assume that he would recognize intentional action as a subset of voluntary action, given that he thinks chosen (προαιρετική) action is such a subset (*NE* 1111b7).

³ Anscombe (1958, §48), and see also Anscombe (1963). Hampshire (1983) defends an even stronger thesis: "To say that I did something intentionally is to say that I knew what I was doing when I did it" (p.145). Other defenders of (PKP) include Newstead (2006), Rödl (2011, 2007), and the authors mentioned in fn.6.

⁴ I will omit the bracketed qualifications from now on unless they are important.

Central to Anscombe's understanding of (PKP) is the contention that the knowledge that agents have of their own actions is of a special kind. This knowledge is 'non-observational' (p.13ff.) and "practical" (p.57ff.) rather than "speculative".⁵ To explain what this means, she uses a formula from Aquinas: unlike speculative knowledge, which "is derived from the objects known", agential knowledge is "the cause of what it understands" (p.87). As we shall see, however, how to understand agential knowledge is a difficult and contested question. The central aim of this work is to present an account of the phenomenon. I begin with Anscombe because she identifies some essential features of agential knowledge that an account of the notion should aim to elucidate. Central among them is the non-accidental connection between agential knowledge and intentional action that (PKP) aims to capture.

For many years there was a consensus in the literature that (PKP) should be rejected in light of counterexamples by Donald Davidson (2001a, 2001b). The force of these counterexamples, however, has recently been contested, as a number of scholars have argued that attention to aspectual distinctions shows that they miss the mark.⁶ As a result, (PKP) has enjoyed something of a resurgence, and it is once again considered an important contender among principles that might elucidate the nature of agential knowledge.

In this chapter I present a novel objection to (PKP), arguing that the thesis is subject to a form of Timothy Williamson's influential anti-luminosity argument (2000). Together with other plausible principles about the nature of knowledge, (PKP) leads to contradiction

⁵ See Moran (2004) on the relation between these two characterizations.

⁶ See Small (2012), Stathopoulos (2016), Thompson (2011), and Wolfson (2012). Thompson's second chapter in *Life and Action* (2008), circulated for a long time before publication under the title 'Naive Action Theory', seems to be largely responsible for the resurgence of interest about aspectual differences in much recent philosophy of action. To my knowledge, Falvey (2000) was the first to point out the importance of aspect for studying agential knowledge, but unlike recent defenders of (PKP) he seems to think Davidson's counterexamples should lead us to weaken the principle (see pp.36-7).

and should therefore be rejected. Since the argument relies on general epistemic principles, rather than on intuitions about fringe cases, the recent responses that have been given to defuse the force of Davidson's purported counterexamples are silent against it. The upshot, if I am right, is that the characteristic properties of intentional action cannot ensure that our beliefs are epistemically safe so as to constitute knowledge. Hence, we shall need a different account of the non-accidental connection between intentional action and agential knowledge than defenders of (PKP) have proposed. I close by suggesting that the connection might be captured more realistically if we take perceptual knowledge as our model for understanding agential knowledge. This suggestion becomes one of the guiding threads for the positive account of agential knowledge presented in Chapter 5.

The chapter is organized as follows. I first explore the reasons to hold (PKP), paying particular attention to an argument on the basis of aspectual considerations. I conclude that the standard objections in terms of presumed counterexamples leave us at an impasse (Section 1). I then present the anti-luminosity argument as it applies to practical entities such as intentional action and intention, leading to the conclusion that (PKP) and related theses drawing a necessary connection between practical entities and knowledge should be rejected (Section 2). I end with a positive proposal for how to capture the central motivations for (PKP) without falling prey to anti-luminosity arguments (Section 3).

1. (PKP) AND PRESUMED COUNTEREXAMPLES

(PKP) expresses an ostensibly necessary connection between intentional action and knowledge. Related theses about other practical phenomena, like intentions, have been

defended by other philosophers.⁷ For now, though, I want to focus on (PKP), in light of its impressive history and given the recent attention it has received.

A few remarks about the principle are in order. First, as the parenthetical phrases indicate, the description under which the agent knows must match the description under which the action is intentional. Second, I shall assume (for reasons that will soon become clear) that (PKP) is aspect-sensitive: it is a thesis about what an agent *is doing* (imperfective aspect), rather than what she *will have done* (perfective aspect). Third, since it is often possible to distinguish between Φ ing *intentionally* and Φ ing *simpliciter*, we might ask whether simply knowing that one is Φ ing *simpliciter* is sufficient to satisfy the knowledge requirement in (PKP). Although it is not always made clear, I assume supporters of (PKP) think only knowledge of an action as done intentionally fulfills the principle, since it would be strange to think that my agential knowledge could be expressed by saying, for instance, ‘I’m walking, but I’m not sure whether I’m doing it intentionally’. Moreover, since (PKP) includes in its scope actions that, as Anscombe notes (1958, §47), must be carried out intentionally (her examples include calling, greeting, signalling, and marrying), an interpretation along these lines is most principled, keeping the knowledge requirement consistent through all action types.⁸

⁷ For instance, Fleming (1964) defends the view that intention requires knowledge that one intends: “to intend to do *x* is itself to know that you intend to do *x*: intending is itself awareness of intending” (p.315). And, as I’ll note shortly, although Davidson is often portrayed as the main opponent to Anscombe’s view, he held a related thesis, holding that the agent must know what she is doing under at least one description if her doing is intentional. Searle (1983) defends a similar view, holding that “at any given point in a man’s conscious life he knows without observation the answer to the question ‘what are you now doing?’”, where knowledge that one is trying to act a certain way counts as meeting this condition (p.90). I shall show that the anti-luminosity argument also calls these weaker views into question.

⁸ These are the actions denoted by what Jonathan Bennett calls “intention-drenched verbs” and Michael Moore calls “intentionally complex” (Bennett 1988: 205-6; Moore 2010: 174). The existence of such actions is sufficient for the anti-luminosity objection I present to (PKP), since the argument below would show that the principle is false at least with respect to these actions.

One reason (PKP) seems attractive is that, as Anscombe notes, if you asked a man why he was sawing a plank (as he moved a saw in the relevant way across a plank) it would be strange for him to reply that he didn't know he was sawing a plank (1958, §6). Such a reply would call into question whether he was sawing the plank intentionally, suggesting that knowledge of the relevant sort is necessary to act intentionally.

Here, however, I want to focus on an argument for the principle that specifically appeals to the progressive nature of agential knowledge.⁹ Thus, take a sentence with a telic verbal phrase with imperfective aspect, such as:¹⁰

(1) The stone is rolling down to the bottom of the hill.

(1) can be true, even if it also turns out that:

(2) The stone didn't roll down to the bottom of the hill.

This can happen, for instance, if the ball was rolling, but someone stuck her hand out and stopped it before it could reach the bottom. Generalizing and ascending semantically: something can be engaged in a process even though the process is never successfully accomplished.¹¹ This feature is known as the 'openness' of the imperfective.¹²

As recent defenders of (PKP) emphasize, we should not conclude from this that the truth-conditions of corresponding imperfective and perfective sentences are wholly

⁹ A version of this argument is presented by Small (2012) and Wolfson (2012). My rendering of it abstracts from differences in their presentation (differences of importance are noted in the footnotes).

¹⁰ I follow defenders of (PKP) in focusing on telic VPs (commonly referred to as 'accomplishments', following Vendler (1957)), and the recent linguistic literature that treats aspect as a property of VPs rather than verbs. See Rothstein (2004) for a good summary of the recent linguistic literature on aspect.

¹¹ I use 'process' in what I take as a now standard way in philosophy of action to refer to those entities subject to predications that differ in aspectual character (see e.g. Steward (1997, 2012) and Stout (1997))

¹² The term comes from Falvey (2000: 22). It is also known to linguists as the 'imperfective paradox', but as Zoltán Gendler Szabó notes, this is a misnomer: "The imperfective paradox is not really a paradox, just a refutation of the once common view that lexical aspect can be analyzed with the resources of classical tense logic" (2008: 523n33)

independent of each other.¹³ After all, if I flick the stone in question from the top of the hill with just enough force to move it a few centimeters downward, that would not be enough to guarantee the truth of (1). Intuitively, this is because in order to be rolling down to the bottom of the hill, the rock must be engaged in a process where the default is for the rock to reach the bottom of the hill: if, in the end, it does not reach the bottom, something must have interfered *preventing* it from doing so. Generalizing, then, the proposal is that:

(3) x is Φ ing iff x is doing something such that it will not be accidental if x will have Φ 'd (i.e. it will have Φ 'd unless something interferes).

The principle expressed by (3) is a general proposal about the truth conditions of telic imperfective sentences (or, ascending semantically, about the nature of telic processes in general). The specific case of intentional action is special, however, insofar as the aim of the action, and the non-accidental connection between it and what the agent does, obtain in virtue of her own representation of what she is doing.¹⁴ Thus, for an agent to be Φ ing intentionally it must be that: (i) she represents what she is doing as being such that it will be non-accidental if she will have Φ 'd, and (ii) the non-accidental connection obtains in virtue of the agent's practical representation of it—a plan or instrumental order which determines what counts as an interference.¹⁵ An action of Φ ing is intentional only if it occurs as an expression of the agent's know-how or skill to carry out that plan; without

¹³ See Small (2012: 179-80) and Wolfson (2012: 331-3).

¹⁴ See Small (2012: 163 et passim) and Wolfson (2012: 344-5) who quotes approvingly the following passage from Searle (1984: 58): “what I am doing depends in large part on what I think I am doing”. This is also a central thesis in Rödl (2007), who identifies a certain form of so thinking with the intentional act.

¹⁵ Small (2012) appeals to the notion of an instrumental order, as characterized by Anscombe's (1958) special sense of the question ‘Why?’, whereas Wolfson (2012) appeals to Bratman's (1999 [1987]) notion of a plan.

the know-how the connection between the representation and the happening will be accidental, conflicting with (i).

Why think the non-accidentality must be this strong, rather than hold that it suffices to act intentionally if the agent accomplishes an aim she ‘sets for herself’? To answer this challenge, Wolfson (2012: 349) appeals to deviant-chain cases, where an intentional action fails to obtain despite meeting the weaker condition. Thus, consider the following case from Davidson (1973):¹⁶ If a man wants to kill someone by shooting a gun, it does not follow that every death that results from the shooting of his gun is intentional. For instance, if a person dies from a heart attack after hearing the shot that the man aims at another person, the killing of the one who dies from a heart attack was not intentional. According to Wolfson, the reason is that the death wasn’t brought about as an execution of his plan.¹⁷

It might at this point help to illustrate the two conditions with an example. Thus, suppose Ann is walking to the store (intentionally). It must then be that: (i) Ann represents what she is doing (e.g. taking some steps) as being such that it would be no accident if she gets to the store as a result, and (ii) the fact that it wouldn’t be an accident is explained by her representation of a plan to get to the store (she is taking these steps to reach the corner, in order to turn left, in order to keep walking to eventually get to the store), a plan that she is executing as an expression of her know-how. An action is intentional, on the proposed view, only under those descriptions that meet these conditions.

Let c be an arbitrary case where S is Φ ing. Given (i), S must have a representation that she will Φ unless interrupted. Hence, she must have a representation that she is Φ ing,

¹⁶ Attributed to Daniel Bennett.

¹⁷ See Mele and Moser (1994) for a similar explanation.

a representation that either constitutes or puts the agent in a position to form a belief that she is Φ ing.¹⁸ Since S is indeed Φ ing in c , the belief is true. Moreover, given (ii), the agent must possess know-how to ensure that unless something intervenes, the agent will succeed in Φ ing. The possession of this know-how ensures that the agent's belief that she is Φ ing is not just true, but non-accidentally true. Thus, it seems to constitute knowledge.¹⁹ Hence, an agent cannot act intentionally unless she knows what she is doing. Metaphysical considerations about the structure of processes therefore give us strong reasons to accept (PKP).

On the other hand, (PKP) was near-unanimously rejected for many years by philosophers of action, who took it to be vulnerable to counterexamples. Here is a famous one by Davidson:²⁰

[I]n writing heavily on this page I may be intending to produce ten legible carbon copies. I do not know, or believe with any confidence, that I am succeeding. But if I am producing ten legible carbon copies, *I am certainly doing it intentionally.* (2001b: 97)

Davidson is arguing here against a weaker thesis than (PKP), namely, the view that acting with an intention requires a belief that one is doing what one intends.²¹ But since knowledge entails belief, the example, if successful, would show that (PKP) is false. More recently,

¹⁸ Small (2012) argues for the first view, whereas Wolfson (2012) seems committed only to the second.

¹⁹ Small (2012: 205-06) moves from the claim that the belief is true and justified to the conclusion that it is knowledge; but, as is well known from Gettier cases, there can be justified beliefs that aren't knowledge. I therefore think the argument is better stated in terms of non-accidentality with respect to the truth. In Section 3, however, I shall argue that the non-accidentality in play is not sufficient for knowledge.

²⁰ What follows is from a passage in "Intending". An earlier and more condensed formulation of the example appears in his "Agency" (2001a: 50). A similar example is offered by Bratman (1999 [1987]: 37).

²¹ And ultimately the target of the objection is the view defended by Grice (1971) and Harman (1986a, 1997, 1986b) that intending to Φ requires belief that one will Φ . See Pears (1985) and Velleman (1989: 114-21) for sustained replies to Davidson's objection.

Sarah Paul (2009a) has argued that cases of absent-mindedness might also be counterexamples to (PKP). For instance, she notes that someone driving on ‘autopilot’ might signal to turn left out of habit without realizing that she is doing so at all (p.5). She suggests that in such a case the person acts intentionally without knowing what she is doing.

Many respond to these cases by rejecting (PKP). Paul, for instance, concludes that there is no interesting necessary connection between practical and cognitive attitudes (2009a, 2009b), defending an inferentialist account of agential knowledge.²² Others try to preserve a connection by weakening the principle. This is Davidson’s preferred strategy. He writes: “Action does require that what the agent does is intentional under some description, and this in turn requires, I think, that *what the agent does is known to him under some description*” (Davidson 2001a, p.50, emphasis added).²³ Similarly, although cognitivists like David Velleman (1989) and Keiran Setiya (2008, 2009) reject (PKP), they aim to capture the insight behind it with principles that appeal to weaker cognitive attitudes than knowledge, beliefs in Velleman’s view, and increases in confidence in Setiya’s (considered) view.²⁴

The third strategy, one that was neglected for many years but has recently received increasing support, is to deny that the presumed counterexamples actually tell against

²² This account is considered at length in Chapter 3.

²³ Falvey (2000: 36-7) also seems to endorse this view.

²⁴ Velleman (1989) initially stated his view as one where intentions are beliefs, but in the introduction to the latest edition of his book, he rejects this formulation (without changing the substance of his view) (2001: xix). What specific beliefs Velleman would identify with intentions is a matter of dispute. See Yaffe (1995) for an elucidative account. Cognitivist accounts like Velleman’s are discussed at length in Chapter 2.

(PKP).²⁵ For instance, it might be claimed that if we allow that knowledge and belief are sometimes unconscious and dispositional states (as opposed to conscious, occurrent states), examples like that of the habitual driver are consistent with (PKP), since the driver might still know what she is doing, though unconsciously and/or dispositionally.²⁶ And while many took Davidson's counterexamples as decisive, the considerations about aspectual differences that ground the metaphysical argument for (PKP) presented above have also been used to call into question the force of the counterexample (Thompson 2011; Small 2012; Wolfson 2012; Stathopoulos 2016). For instance, Michael Thompson (2011) argues that in the ordinary cases of *making* (progressive) ten carbon copies, one has the opportunity to check and try as many times as needed until one has succeeded in one's aim. In those cases, all of what one does is subsumed under the action-description *making ten carbon copies*. Under that description, the agent can know all along what she is doing, even if she fails on her first try (and even if she fails to bring the action to completion). Of a case where the person only gets one chance, Thompson writes:

[F]or him, the making of the inscription is like the buying of a lottery ticket. You can say he made ten copies intentionally if you like, but it will not be an illustration of the topic of Anscombe's book, any more than lottery-winning is when you bought the ticket with that aim. (p.210)

Small (2012) arrives at the same verdict:

[I]t's not clear that a copier under such a strange and unusual demand is in a position to intend to make all ten copies at once. He can aspire to, perhaps . . . he can give it a shot and hope for the best. But if you hope to make ten copies and you bring it about that ten copies are made, you haven't necessarily acted intentionally. (p.199)

²⁵ For an early precursor of this type of response, which distinguishes between the sense of 'intentional' of interest to philosophy of action and the sense employed in ordinary language, see Velleman (1989: 114-21).

²⁶ See Setiya (2009: 389) for a response along these lines. A related strategy is to hold that the person acting need not *actually* know what she is doing, but must at least be in a position to know (see fn.29 below on 'being in a position to know'). As an anonymous referee pointed out, though, these responses are complicated by the fact that, as Paul imagines the case, the person can come to know only *by looking* at the dashboard.

Indeed, as both Small (2012) and Wolfson (2012) note, their metaphysical argument for (PKP) entails that under such conditions the carbon-copier is not making ten carbon copies intentionally, since the action's success is in this case overly accidental.

Now, I don't take these authors to deny that ordinary intuitions about the case might lead one to say that the man is intentionally making ten carbon copies in the case where he only gets one go.²⁷ Rather, their point is that such an example is very different from the central cases of what we would subsume under the concept INTENTIONAL ACTION where, in particular, the person retains control over her action in a way that makes her success non-accidental: the person has the chance to correct and continue pursuing her intended action in the face of obstacles and difficulties. Therefore, a theory of intentional action need not be responsive to our intuitions about these cases.²⁸

Drawing a parallel to a dispute in epistemology is helpful at this point. Many of us hold that knowledge entails belief: Necessarily, if *S* knows *p*, *S* believes *p*. However, there are cases where some find it intuitive to ascribe knowledge while denying belief.²⁹ Although this leads some epistemologists to reject the view that knowledge entails belief, many of us judge that given how much else there is to be said for it, we should hold on to the connection even if it conflicts with some intuitions: our theory of knowledge is allowed

²⁷ The studies presented in Chapter 2 suggest that people generally think in cases with this structure a person acts intentionally and without knowledge. I should note, however, that I do not readily share the intuitions myself; but to the extent that we can ascribe evidential value to intuitions, these ones seem safe enough.

²⁸ Wolfson (2012) is particularly clear on this score when he says that the conception he defends implies that "we do fewer things intentionally than many philosophers, and many non-philosophers, have thought" (p.354).

²⁹ The contention goes back to Radford (1966), and has been recently defended by Myers-Schulz and Schwitzgebel (2013).

to diverge from fringe intuitions. What goes for the connection between knowledge and belief should go for the connection between knowledge and action. Given the arguments that motivate defenders of (PKP), the lack of control in the carbon-copier case gives us reason to think it is fringe. Hence, those attracted to the principle are free to reject objections based on intuitions about such cases.

2. PRACTICAL ANTI-LUMINOSITY

There are powerful arguments in favour of (PKP), and outstanding objections to it are not dialectically strong. In this section I shall argue that, nonetheless, we should reject it because it conflicts with principles that capture important properties of knowledge. This is shown by the fact that (PKP) is subject to arguments of the sort that Williamson (2000) has used to call into question the luminosity of mental states like appearances, feelings of pain, and knowledge.³⁰ Because it is not reliant on intuitions about fringe cases, this strategy cannot be dismissed on the same grounds as Davidson's counterexample.

Let's start with a brief summary of the anti-luminosity argument for the case Williamson considers, that of feeling cold (a more detailed version is given below for practical cases). Consider the following thesis:

Luminosity of cold feelings (LUM):³¹ Someone feels cold iff they are in a position to know that they feel cold.

³⁰ A condition *C* is luminous just in case, the following is necessarily true: one is in *C* iff one is in a position to know that one is in *C*. To be in a position to know *p* is to be such that if one is doing what one can with the aim of finding out whether *p* (such as directing one's attention in the relevant way), then necessarily one knows *p*.

³¹ There is no standard labelling of the principles in the anti-luminosity argument in the literature. I largely follow Srinivasan's (2015).

Williamson's argument against (LUM) begins with the observation that we often go from feeling cold to feeling warm slowly and gradually. Suppose that such a transition takes place as the sun warms up Sunny in an initially cold morning. As time goes by, it becomes less and less clear whether Sunny is feeling cold in such a case, and Sunny will start losing confidence that she is so feeling as a result. At some point, Sunny's confidence will be fairly low but still high enough for belief. In that case, there are two possibilities: either her belief that she feels cold is true or it isn't. If it isn't, then clearly Sunny doesn't know; but even if the belief is true, Sunny will not know because there are nearby cases where her belief that she is feeling cold is false: her belief is too accidentally true to constitute knowledge.

A number of strategies have been proposed to resist the anti-luminosity argument.³² It is beyond the scope of this paper to consider them. My view is that the argument is sound.³³ But it is sufficient for my purposes that many philosophers endorse the arguments and that even its detractors admit that there are important costs to rejecting it.³⁴

It is easy to adapt the argument about feeling cold to show that many other conditions are not luminous. What has not been noticed, and I want to argue here, is that it can be adapted to practical phenomena like intention and intentional action. Anti-

³² See e.g. Berker (2008); Brueckner and Fiocco (2002); Cresto (2012); Egré and Dokic (2008); Egré (2006); Leitgeb (2002); Neta and Rohrbaugh (2004); Vogel (2010); Weatherson (2004); Cohen (2010); Blackson (2007).

³³ Some of the central objections are, in my view, successfully addressed by Williamson (2005) and Srinivasan (2015), and the considerations presented below in defense of the application to the practical cases can be used as a basis to defend the argument against some of the most powerful objections.

³⁴ Thus, surveying iteration principles of the sort that the anti-luminosity argument calls into question, Daniel Greco, who has elsewhere defended the luminosity of knowledge (Greco 2014), notes that rejection of margin for error principles "will be a bullet to bite" (2015: 769). In a similar vein, Jonathan Vogel notes that there is "something mysterious about the luminist's position" given the commitments she incurs when rejecting the argument (2010: 564).

luminosity arguments present a powerful challenge to (PKP) and related theses, a challenge that I believe should lead us to reject these principles.

Consider the following case:

Cleaning Sisyphus: The floors of the Underworld are getting filthy with blood and bile. Hades notices, and decides to give Sisyphus a more useful task than tossing a rock around. Handing him a mop, he orders him to clean the floors using the waters from the river Acheron, the cleanest in the Underworld. There is one problem: while at noon the waters of Acheron are clean as a spring, they slowly and gradually get dirtier and dirtier as the day goes by—by midnight it is just filth, much dirtier than the floors of the Underworld. Hades thus tells Sisyphus that at midnight he will be punished proportionally to his efficiency: his punishment will be worse if either he fails to keep mopping when the water is still clean enough or if he keeps mopping when the water is dirtier than the floors. Each day, therefore, Sisyphus grabs his mop in the morning trying to mop as much of the Underworld as he can while the water is still cleaner than it. At noon, he is fully confident that he is intentionally acting under the description *cleaning the floors*. As the day goes by, however, he loses more and more confidence. By midnight, even in those days when he keeps mopping to defy Hades, he is certain that he is *not* cleaning the floors of the Underworld (but instead making them dirtier).

We can assume that Sisyphus is very skilled at mopping floors, and that he knows this. This ensures that Sisyphus meets the know-how condition that defenders of (PKP) argue is necessary to engage in intentional action. With this in mind, take the span of a day where Sisyphus mops continually from noon to midnight, concentrating throughout on what he is doing.³⁵ Let t_0, t_1, \dots, t_n be a series of times at one millisecond intervals from the time he starts mopping until a time where he is still mopping but no longer cleaning the floors because the water is too dirty, and let α_i be the case at time t_i in this series.³⁶ By the description of the case, at α_0 Sisyphus is cleaning the floors (intentionally), and knows it,

³⁵ The last supposition is needed to ensure the transition from being in a position to know to knowing. As Antonia Peacocke suggested, it is possible that the transition is different for practical and theoretical cognition, in ways that might complicate the argument (e.g. since paying attention in the practical case will increase the chances of success). The case of Cleaning Sisyphus is set up to fend off these worries (throughout, he is attending to his task).

³⁶ A case is a centered possible world.

while at α_n he isn't cleaning the floors, and knows it. It seems plausible to think that Sisyphus can't discriminate between cases α_i and α_{i+1} with respect to what he is doing. The following principle therefore seems to apply (more on this soon):

Margin for Error for Action (MARA): If Sisyphus knows that he is cleaning the floors at α_i , he is cleaning the floors at α_{i+1} , for all times in the series t_0, t_1, \dots, t_n . From (MARA), it follows that Sisyphus is cleaning the floors at α_1 , given that he knows he is doing so at α_0 . Now suppose (PKP) is true. If so, Sisyphus knows that he is cleaning the floors at α_1 , and again, by (MARA), that he is doing so at α_2 ; hence, by (PKP), he knows he is cleaning the floors at α_3, \dots and so on. Evidently, sufficient applications of this mode of reasoning (n-many) will yield the conclusion that Sisyphus is cleaning the floors of the Underworld at α_n . This contradicts the initial assumption that he isn't doing so at the time. (PKP) leads to contradiction, so it must be rejected.

As I said above, a number of strategies have been proposed to resist the anti-luminosity argument, and it is not my aim to defend it against them here. However, I will now consider objections that might arise from the specific nature of intentional action and agential knowledge that might not generalize to applications of the anti-luminosity argument to other conditions. Before doing so, though, I want to briefly defend (MARA), given the crucial role that it plays in the argument.

(MARA) can be seen as a way of capturing, in modal terms, the view that to constitute knowledge, a belief must be non-accidentally true.³⁷ This property becomes

³⁷ Williamson defends margin for error principles by appeal to a safety principle, according to which knowledge requires it to be the case that in nearby cases where one forms a similar belief, the belief is true. Other defenders of the safety principles include Ernest Sosa (2000, 1999) and Duncan Pritchard (2007, 2008, 2009). A proper defence of the principle would require defending the transition from safety to margin for error (for which see Srinivasan (2015)).

salient when we consider standard fake-barn cases.³⁸ Suppose I find myself in a field full of objects, most of which are fake barns, and only one of which is a real barn. I point to one of them and say, ‘That’s a barn’. It seems that even if I am pointing to the real barn I could not know that what’s in front of me is a barn. Here is an attractive explanation of what goes wrong: given the presence of all those fake barns nearby, there are nearby cases where a similarly formed belief would be false. By contrast, if I was in a similar position but the field was full of real barns instead, there would be no nearby cases where my belief that there is a barn would be false. In such a case, then, my belief would be non-accidentally true, and I would know. This explanation relies on something like a generalized version of (MARA)’s being true.

Recent defences of (PKP) appeal to a distinction between action in the progressive and completed action. A central difference between them, we saw, is that it is possible to *be doing something* even when one will go on to fail. It might thus be suggested that in the example, Sisyphus *is* still cleaning the floors (if he so intends), even when the water is pure filth; it’s just that he’s not succeeding at achieving his goal. However, for this strategy to work as a defense of (PKP), we would need to hold that one can be engaged in an action *regardless* of the material conditions in one’s environment, which is utterly implausible: I can’t be crossing a road in a desert (regardless of my intentions) if the closest road is many

³⁸ The cases were popularized by Goldman (1976), and are credited to Carl Ginet. Recent x-phi work in this area suggests that the intuitions in these cases are not as straightforward and secure as it might have been thought initially (Colaço et al. 2014; Horvath and Wiegmann 2015). With many others, however, I think there are good reasons to hold that fake-barns subjects lack knowledge. Moreover, to draw out the right intuition, it helps to imagine the case in such a way that the person has consistently gotten matters wrong (DeRose 1995: 30). This last insight, I have learnt, is owed to Helen Beebee.

kilometers away.³⁹ By the same token, regardless of his intentions, Sisyphus can't be cleaning the floors if the water is so grimy that it is instead clearly dirtying up the floor. If my friend told me that she is cleaning the floors of her apartment using dungy mud, I have to think that she's making a complex joke, or that she has gone mad.

This, it should be noted, is a point that is emphasized by defenders of (PKP) who, following Anscombe, take the principle to be about something that 'happens' rather than an inward mental state.⁴⁰ This is part of what makes (PKP) such an interesting thesis. However, the point shows that what I consider the most promising response to the anti-luminosity argument cannot easily be adapted as a defence of (PKP).⁴¹ The response hinges on the contention that there are beliefs about certain facts, call them 'B-facts', that meet the following condition:

(*) If S believes that a B-fact obtains, a B-fact obtains.

But if (*) holds, then we could imagine a case where S has the belief that a B-fact obtains even though in nearby cases the belief would be false. If so, there are beliefs about B-facts that don't meet a margin for error principle *but* nonetheless count as knowledge because the fact that they are true is not a matter of luck: (*) ensures that these beliefs are true whenever held. Now, it is disputed whether this strategy works;⁴² but the point I want to stress here, one emphasized by those who endorse this objection, is that the only reason it

³⁹ The possibility of error here grounds the fact that, as Donnellan (1963: 403) points out, one can make assertions of the form 'I thought I was Φ ing, but in fact I am not \neg ', e.g. 'I thought I was turning on the radio, but I'm not since the radio doesn't work'.

⁴⁰ See Wolfson (2012: 329, 37) and Small (2012: 167-73).

⁴¹ This type of response was first presented by Leitgeb (2002) and is developed by Weatherson (2004) and Berker (2008).

⁴² For the record, I do not think they work, for the kinds of reasons presented by Srinivasan (2015).

is plausible to think that beliefs about B-facts might constitute knowledge despite not meeting a margin for error principle is that a condition like (*) holds for them. However, we have just seen that such a principle cannot hold for intentional action given its dependence on material conditions.⁴³ Anti-luminosity arguments are thus particularly powerful as objections to (PKP).⁴⁴

On the other hand, I want to consider two objections to the principle that appeal to what's distinctive of agential knowledge as "the cause of what it understands", to call into question important assumptions in the argument as given. Let's distinguish two objections.⁴⁵

The first objection is that it cannot be assumed that a situation analogous to a fake barns case could arise for agential knowledge. The reason is that whereas things like barns (both fake and true) exist independently of what the agent thinks about them, intentional actions are dependent on the agent's thought for their existence. This contention, as we saw, is a central motivation for recent defenders of (PKP), who hold that whether one is Φ ing intentionally depends precisely on whether one represents what one is doing as a

⁴³ More carefully: the occurrence of (at least some) intentional actions that lie in the scope of (PKP) depend on material conditions extrinsic the agent's state of mind. Of course, on a materialist understanding of the mind every state of mind depends on the occurrence of certain material conditions; but these conditions will be intrinsic to the state of mind. It is also important that I am presently concerned with (PKP) in an unrestricted form, applying to all intentional actions, and only later will consider how anti-luminosity arguments call into question even more modest principles. For all I say here, a strategy appealing to B-facts might be viable as a defense of a restricted version of (PKP), say, one restricted to tryings understood à la Hornsby (1980). Note, however, that such a strategy would have to show how it gets around the arguments below that suggest that the meeting a condition like (*) is not enough for knowledge, if the source of the belief is itself unreliable. Thanks to the editors from MIND for encouraging me to clarify this point.

⁴⁴ Likewise, strategies that deny the applicability of safety (and therefore margin for error) principles to second-order (but not to first-order) knowledge also fail as defenses of (PKP) assuming actions are not mental states (Egré and Dokic 2008; Egré 2006; Cresto 2012).

⁴⁵ I thank two anonymous reviewers for these objections.

Φ ing. For this reason, it is unclear whether an agent could ever find herself in a situation where she believes that she is Φ ing intentionally, is doing so, but there are nearby cases in which she is not.

A related objection directly questions whether a margin for error principle applies to agential knowledge. For it may be suggested that part of what it is for such knowledge to be “the cause of what it understands” is that in the situations where an agent is in a position to Φ intentionally, her intention settles whether she is Φ ing or not. This is because in these situations she both has the relevant know-how to carry out the action, and the circumstances are such that she could carry it out (since she is in a position to do so). Hence, if the agent’s beliefs are constituted by or aptly track those intentions, the beliefs will necessarily be true. In the case of Sisyphus, for instance, there will come to be an interval of time (say, t_i), at which he is cleaning the floors intentionally (and so believes), but at the end of which he ceases to intend to clean the floors because the water is getting too dirty. He would thereby cease to intentionally clean the floors and to believe that he is doing so at t_{i+1} . However, the fact that he is no longer cleaning the floors intentionally at t_{i+1} doesn’t count against the claim that his belief constitutes knowledge at t_i , since he is capable of discriminating between these cases through his *practical* discernment: his intention to cease cleaning is enough to distinguish the case where he is intentionally cleaning and where he is not. Therefore, his beliefs cannot but be true when they are aptly formed on the basis of his practical capacities. This means that, despite not meeting (MARA), these beliefs are sufficiently reliable to constitute knowledge in the case where the agent acts intentionally.

Let's take the objections in turn. Starting with the first objection, I shall aim to show that the same kinds of considerations about the dependency of intentional action on material conditions show that one can find oneself in epistemically perilous cases, analogous to a fake-barns case. To see this, consider the example Anscombe introduces in *Intention* to elucidate the notion of practical knowledge. She writes:

Imagine someone directing a project, like the erection of a building which he cannot see and does not get reports on, purely by giving orders. His imagination (evidently a superhuman one) takes the place of the perception that would ordinarily be employed by the director of such a project. He is not like a man merely considering speculatively how a thing might be done; such a man can leave many points unsettled, but this man must settle everything in a right order. His knowledge of what is done is practical knowledge. (p.82)

The point of considering such a weird case is clear: Anscombe is seeking to isolate a distinctively practical source by which the director might know what is happening with the project, so she needs to ensure that in the example the director does not know of this through theoretical sources, like reports or perception. How, then, *does* he know? He knows on the basis of his orders: so long as they are executed at some given stage, and his imagination represents them as being followed as they are, he can continue giving orders until the project is finished. And if the workers are obedient, and the director's imagination is working properly, he will have knowledge on the basis of his orders that a construction project of a determinate sort (say, a house with two floors, one bathroom, etc.) is being carried out; and since the source of the knowledge is practical rather than theoretical, then, as Anscombe writes, "His knowledge of what is done is practical knowledge".

The terms of the analogy are also clear. As the director stands to the project, an agent stands to her intentional actions. In particular, as the orders of the director determine, in a distinctively authoritative way, what gets done in the project, an agent's intentions

determine, in a distinctively authoritative way, what she does intentionally. And as the director knows of the project through his orders, an agent knows what she does through her intentions.

Imagine now that the case is as Anscombe presents it, with the director giving orders at a distance without theoretical checks on what happens. Today, the director is giving orders to the electrician, who, luckily, is perfectly obedient and skilled at her job. His orders to her, therefore, are executed perfectly, so the director's beliefs about what she does are true. Imagine, however, that all the other workers are either completely disobedient or incompetent, so that when the director orders them to do something, his orders are never fulfilled. I submit that under such circumstances, the beliefs that the director forms on the basis of his orders to the electrician are analogous to the beliefs that the person forms in a fake-barns case: though the beliefs are true, there are too many nearby cases where similar beliefs are false. Hence, they don't amount to knowledge.⁴⁶ This case, moreover, helps to diagnose *why* such cases can arise for agential knowledge, and, therefore, why anti-luminosity arguments are particularly powerful against (PKP): even granting that agents have a practical and authoritative kind of knowledge of their actions, limited agents like us do not have such knowledge of the material conditions of our actions: the director does not have practical knowledge of whether his workers are obedient or competent, and Sisyphus does not have practical knowledge of the conditions of cleanliness of the water. These limitations make perilous epistemic situations inescapable for agents like us.

⁴⁶ Working independently, John Schwenkler (2015: 22-25) arrives at a similar conclusion on the basis of these cases.

Moving on to the second objection, it is worth noting that Williamson (2000) offers a second version of the anti-luminosity argument that aims to target precisely conditions like B-facts; and such a version of the argument would still apply to the case of agential knowledge. I cannot consider the argument in detail, but I'll present a brief version.⁴⁷ The basic idea is to run a parallel anti-luminosity argument targeting, not the person's beliefs, but her epistemic confidence. When one forms beliefs about a condition that changes slowly and gradually, one's confidence in the proposition that the condition obtains slowly diminishes as it becomes less and less clear that the condition obtains. As we reach an epistemically perilous case, a person's confidence level will either be high enough to constitute knowledge, or it will not. If it is not, then she doesn't know; but if it is high enough, then her confidence will be out of proportion to what the case merits (given that she is in an epistemically perilous situation), and her mental state will therefore be too unreliable to constitute knowledge. Either way, then, the person will fail to know, and all of this is consistent with the supposition that the agent forms only true beliefs about the matter at hand when these beliefs track her intentions.

Appealing to the type of considerations that showed that epistemically perilous situations arise for agential knowledge, however, we can give a more direct response to the second objection.⁴⁸ To do so, we need to take a deeper look at margin for error principles. As noted, these principles are part of a larger class of modal principles that aim to capture the characteristic non-accidentality of the relation between truth and belief when one has

⁴⁷ I refer the reader to Srinivasan (2015, p.15ff.) for a detailed account and defense of the argument.

⁴⁸ Because the objections above have a structure similar to the one pursued by Berker (2008), they provide the bases for a more general defense of the anti-luminosity argument against such objections. I hope to explore the point further in future work.

knowledge. It is generally agreed that these principles are *source-relative*: to decide whether a possibility is epistemically near, we can ignore worlds where the agent would form false beliefs on the basis of a different source. Thus, to take a case from Nozick (1981), suppose a woman judges that her grandson is healthy upon seeing him. The knowledge that the woman thereby acquires of her grandson's state would not be called into question even if she would have later judged that he was ill had he called her to say so. Knowledge acquired on the basis of perception is not called into question by the possibility of forming a false belief on the basis of testimony.

However, it is equally clear that the principles cannot be so strict as to require that only beliefs formed on the basis of the *same* source count as nearby. Cases where the person forms beliefs on the basis of a *similar* source must also count as near. To see why, suppose that whenever one *remembers* something, one recalls a fact (i.e. a true proposition). This would ensure that beliefs formed on the basis of remembering would be non-accidentally true, since every such belief will be true. However, suppose that a man suffers from memory-delusions: very often, when it seems to him that he is remembering something, he is actually just making something up. Those delusions do not count as remembrances, since they are not true. However, the beliefs that the man forms on the basis of his remembrances are clearly not reliable in a case like this, given how easily he could be suffering from a delusion when he thinks he is remembering. Therefore, the beliefs that this man forms on the basis of his memory do not constitute knowledge.

Let's now consider the practical case. Grant, for the sake of argument, that whenever a person intends to Φ in circumstances where she is in a position to Φ , she forms a true belief that she is Φ ing. Thus, practical beliefs formed on the basis of intentions are

non-accidentally true, since every belief formed on *that* basis is true. However, it is a standard view in philosophy of action (and one shared by defenders of (PKP)) that there are similar cases where, if one formed the belief that one is acting intentionally, the belief would not be true, because in such cases one's attitude doesn't amount to an intention. For instance, Small (2012: 167-73) follows Baier (1970) in holding that when the action is not appropriately within the agent's control, the agent merely has an aspiration, which falls short of an intention; but—and here is the crucial point—a belief that one is acting intentionally, when formed on the basis of an aspiration would not be true, unlike beliefs formed on the basis of intentions; yet it might be hard for an agent to tell whether what one has is an aspiration or an intention.⁴⁹ An agent might thus find herself in a situation where she cannot tell whether she is intending to Φ , or merely aspiring to do so (hence whether she is acting intentionally). In that kind of case, even when the agent forms a belief on the basis of an intention that ensures that she is acting intentionally, her belief might not constitute knowledge because there is a nearby case where she merely has an aspiration, so that she is not acting intentionally (perhaps she acts 'aspirationally'). At some point in the series, the case of Sisyphus has this structure, and this explains why his beliefs about what he is doing, even when true, fail to constitute knowledge.

I take the above to sufficiently show that anti-luminosity arguments present a powerful challenge to (PKP), one that cannot be fend off by the same strategies as have been employed in response to Davidson's carbon-copier case. Interestingly, these

⁴⁹ As Small puts it (p.170), "there need be no easy answer" for the question whether one is intending to do something or merely aspiring to do so in a particular case.

arguments also call into question more modest cognitivist principles, such as Davidson's weaker view that:⁵⁰

Weak Practical Knowledge Principle (WPKP): If an agent is Φ ing she knows that she is Φ ing under *at least one* description of what she is doing.

To see this, imagine that a group of scientists give you a drug at noon which, they say, will slowly inhibit your motor system so that in three hours you will be unable to move your fingers at all. Suppose they are telling you the truth and you believe them (e.g. based on evidence from previous experiments). For this study, they blindfold you and instruct you to continually move your finger as you usually would. Since the drug has not affected you at noon, when you set out to move your finger at that time you move it intentionally and know this, but by 5pm, you don't move it and know this (even though the situation appears to yourself exactly as though you were). The case so far has a similar structure to Sisyphus's, and it should be easy to see how the argument would go from here to show that you do not know that you're moving your finger in this case. However, the description *moving a finger* seems like the most basic kind of description under which you might know what you are doing for a case like this.⁵¹ Understanding the principle in such a way that even more basic descriptions will count as fulfilling it risks triviality. After all, a sufficiently weak principle will apply to not just actions, but any happening whatsoever: the principle that necessarily, if some process x is happening a person knows that x is

⁵⁰ The version of the argument that follows is also meant to forestall an objection on the basis of a distinction between achievements and activities, to which Stathopoulos (2016) appeals in his defense of (PKP).

⁵¹ I shall go on to consider what might be viewed as more basic descriptions, such as 'trying to move a finger'. However, note that there are views (e.g. O'Shaughnessy (1980)) on which the trying is a different occurrence that precedes the action. Moreover, we could imagine a case with a patient old enough to know that she is moving her finger, but still unacquainted with the concept of trying. Even supposing that 'trying to move a finger' is a description of her action, this might be a description that the patient does not have recourse to.

happening under *at least one* description will come out true if we allow the following description to fulfill it: *something is happening*.

It may be replied that there is a more basic and non-trivial description under which the person might know what she is doing in a case like this, perhaps *trying to move a finger*, or *intending to move a finger*. But anti-luminosity arguments also show that even principles stated in those terms fail. Thus, suppose that at noon I form the intention to go to a party that I have been invited to: if the party were to start then, I would go. I am firmly committed, and I know it. Unfortunately, the party is not until 9pm and I am at home, with nothing to do. Thus, having formed the intention, I sit on my couch through the afternoon thinking about the party. I am the kind of the person who, the more I reflect on something, the less I want to do it. Thus, as I think about whether I intend to go to the party, my determination begins to fade slowly, so that by 9pm I have lost all resolve to go. By 9pm, I therefore no longer intend to go to the party, and know it. Once again, it should be clear how, on the basis of a margin for error principle we can reach a contradiction from the assumption that intentions are luminous. This suggests that we should reject that assumption.

The foregoing argument relies on the gradeability of commitment. Because commitment is a feature that intentions share with the lot of practical entities that are the subject of study of philosophy of action—entities like decisions, attempts, and basic actions—it is straightforward to run an anti-luminosity argument for them on the basis of the one just presented. Hence, I submit, we have good reason to think that there are no

interesting practical entities that are luminous, and, hence, no interesting necessary connection between such entities and knowledge.⁵²

3. RECONSIDERING THE NON-ACCIDENTAL CONDITION

The considerations in the previous sections show that there are powerful objections to (PKP) that do not rely on intuitions about fringe cases, and I have argued that there are no straightforward ways of responding to the argument. Of course, there might be responses available that I have not considered; in that case, I hope the argument at least puts the onus on defenders of (PKP) to present those responses. Someone following this strategy should also say what costs she is willing to pay, since, as I noted, even opponents of the anti-luminosity argument grant that there are costs to pay in rejecting it. Although I am not optimistic about the prospects of this project, it is worth noting that the results would be of independent interest: it has been thought that the most plausible candidates for luminous conditions are internalist or purely mental ones, whereas (PKP) is a principle governing intentional actions, paradigmatic external entities. A successful defense of (PKP) would show that, contra Williamson, there are far from trivial luminous conditions.

On the other hand, the argument seems to leave us at an unsatisfying place. For the argument for (PKP) from Section 1 suggests that there are general truths about the metaphysics of action that support the principle, whereas I've argued that the anti-luminosity argument shows that general truths about knowledge require us to hold that the principle is false. We seem, therefore, to have reached one of those perennial philosophical

⁵² In the present chapter I am concerned only with cognitivist principles stated in terms of knowledge. In Chapter 2 I build on a suggestion by Greco (2015: 768), and apply the principles to call into question cognitivists principles in terms of belief.

impasse. Against this verdict, I shall presently argue that some of the considerations at play in the previous section also suggest a way of accepting the central contentions in the metaphysical argument without accepting (PKP) and, therefore, to move beyond a mere impasse.

At the most general level, the metaphysical argument for (PKP) is motivated by the following three general claims:

- (1) *Representation and content connection*: If S is Φ ing intentionally she must represent herself as Φ ing intentionally.
- (2) *Non-Accidental*: The match between content and representation in (1) is non-accidental; it is not accidental that if S represents herself as Φ ing she is Φ ing (at least in cases where she is).
- (3) *Representation and non-accidental connection*: (2) holds in virtue of essential features of S 's representation qua intention.

From these claims, it is concluded that an agent cannot be Φ ing intentionally without knowing that she is doing so. For suppose that S is Φ ing intentionally. Given (1), the agent must represent herself as Φ ing intentionally, which means that she either believes that she is (if the representation constitutes her belief) or is in a position to so believe, given the representation. Given that we are considering a situation where she is indeed Φ ing intentionally, this belief is true. Moreover, given (2) and (3), the fact that the belief is true is non-accidental and it is partly in virtue of the agent's particular kind of representation that it is non-accidentally true. Therefore, if an agent is Φ ing intentionally, she must have (or be in a position to have) a non-accidentally true belief that she is Φ ing intentionally, a non-accidental that is grounded in her very representation. From this, it is concluded that the agent must (be in a position to) know that she is Φ ing intentionally when she is, as per (PKP).

It is this last step that I want to challenge by arguing that, even if (1)-(3) are true, it doesn't follow that (PKP) is true. To see why, it is helpful to consider a parallel case in the theoretical realm, the case of the connection between perception and perceptual knowledge. In parallel to (1)-(3), the following seems true of perception:⁵³

(4) *Representation and content connection*: If *S* is perceiving that *p* she must have a representation that *p*.

(5) *Non-accidentality*: The match between content and representation in (4) is non-accidental; it is not accidental that if *S* perceives that *p*, *p* is the case.

(6) *Representation and non-accidentality connection*: (6) holds in virtue of essential features of *S*'s representation qua perceptual representation.

Claim (4) is trivial, since perceiving *is* a representing relation. The grounds for (5) and (6) are essentially the same grounds as motivate (2) and (3): we need these claims to account for deviant-causal chains. Just as a person does not act intentionally unless her action is brought about non-accidentally by an intention, a person does not perceive unless her representation of *p* is brought about non-accidentally by *p*. As in the case of intention, this requires more than a causal connection, as brought out by a famous case by Grice:

It might be that it looked to me as if there were a certain sort of pillar in a certain direction at a certain distance, and there might actually be such a pillar in that place; but if, unknown to me, there were a mirror interposed between me and the pillar, which reflected a numerically different though similar pillar, it would certainly be incorrect to say that I saw the first pillar, and correct to say that I saw the second. (Grice 1988: 69-70)

Grice focuses on seeing, but the point obviously generalizes to the other sense modalities.

Given (4)-(6), does it follow that whenever an agent perceives that *p* she knows, or is in a position to know *p*? It does not. We can see why, once more, by reflecting on fake-

⁵³ For simplicity, I shall assume that the content of perception is propositional.

barns cases. In such cases, the person sees a barn in front of her, but is not in a position to know that there is a barn in front of her. This is because the environmental conditions make the connection between her perceptual belief and the truth too epistemically precarious to amount to knowledge. As I argued in the previous section, analogous cases obtain for agential knowledge. Hence, just as the type of non-accidental connection between mind and world that obtains when an agent perceives does not guarantee that the agent is in a position to acquire perceptual knowledge, the type of non-accidental connection between mind and world that obtains when an agent acts intentionally does not guarantee that the agent is in a position to acquire agential knowledge. Hence, the considerations at play in the argument above not only call (PKP) into question, but suggest a way to grant the central claims that motivate recent defenders of the principle.

On the other hand, I think defenders of (PKP) are right to insist that there is an important link between intentional action and agential knowledge; but we do not need to accept a necessary connection between the two to recognize this point (Gibbons 2010). Indeed, the parallel with the perceptual case suggests ways of characterizing the link without recourse to necessary connections. For instance, one might say that just as the function of perception is to yield knowledge, the function of the will (understood as the capacity to act intentionally) is to yield agential knowledge. Additionally, one might hold that just as perceiving can *on its own* (and without the need of an inference) give us knowledge about our surroundings, exercises of the will can on their own give us agential knowledge. If either of these ideas is correct, we would expect that cases where an agent acts intentionally with agential knowledge would form a central core of cases of intentional action. However, since powers can sometimes fail to achieve their function, and since even

non-inferential sources can fail to yield knowledge in epistemically inhospitable circumstances, both ideas are compatible with rejection of (PKP).

Of course, these ideas are in need of defense, and we should not expect an exact parallel, given that perceptual knowledge is paradigmatically ‘speculative’ rather than practical. In Chapters 4 and 5 I attempt to provide a defense of these ideas.⁵⁴ Before doing so, however, we need to consider other approaches that have emerged in the literature on agential knowledge. In the next Chapter, I turn my attention to cognitivist views that attempt to preserve the connection between intentional action and agential knowledge by weakening the cognitive component in a principle like (PKP).

⁵⁴ See Peacocke (2003: 121-23) and O'Brien (2007, ch.9) for further proposals along these lines.

Chapter 2: Against (Traditional) Cognitivism

If, as the argument in the previous chapter suggests, one can act intentionally without knowing what one is doing, then what is the connection between intentional action and practical knowledge? One dominant answer to this question is what I shall call ‘traditional cognitivism’.⁵⁵ Like Anscombe, traditional cognitivists hold that there is a necessary connection between practical and cognitive entities. However, they also hold that the cognitive entities in question are weaker than knowledge. For instance, many endorse the thesis that:

Intention Entails Belief (IEB): (Necessarily) If an agent intends to Φ , she believes that she will Φ .

More recently, Setiya has defended the even weaker view that:

Intentional Action Raises Confidence (IARC): (Necessarily) If S is Φ ing intentionally, S believes that he is Φ ing or is more confident of this than he would otherwise be, or else he is Φ ing by doing other things for which that condition holds.⁵⁶

Are these principles true? If the argument in the previous chapter is successful, it establishes that:

Action Without Knowledge (AWK): It is possible for an agent to Φ intentionally without knowing that she is Φ ing.

⁵⁵ I have picked this label explicitly to distinguish the category from what Paul (2009a) calls ‘weak cognitivism’. On Paul’s usage, a strong cognitivist is someone who holds that there is an *identity* between practical and cognitivist entities, whereas a weak cognitivist merely holds that there is a necessary connection (perhaps explained by a relation of constitution). In Chapter 5 I defend what I call a ‘moderate cognitivism’. Until then, however, I may use ‘cognitivism’ to refer to the traditional version specifically.

⁵⁶ Apart from changes in symbolization, this is a verbatim version the principle defended by Setiya (2008, 2009).

This result can be used to mount an objection to (IEB).⁵⁷ To do so, we need one further assumption:

Knowledge Norm of Belief (KNB): Believe that p only if you know that p .

Why hold this norm? The grounds are similar to those that support the widely endorsed norm for assertion:

Knowledge Norm of Assertion (KNA): Assert that p only if you know that p .

I shall provide these grounds in the next chapter. For now, let us take (KNB) for granted.

Consider now a case where an agent, Ann, intentionally does something without forming an *antecedent* intention, e.g. she scratches her neck. (I am not assuming here the contested view (sometimes ascribed to Anscombe) that one can act intentionally without having an intention at all; rather, I am assuming that in this case, the only intention is an intention in action.) (AWK) was shown to be true for any action type, so we may assume that if Ann scratches her neck intentionally, but she doesn't know she is doing so. Given (KNB), it follows that if Ann believes that she is scratching her neck she believes incorrectly. However, the cognitivist principle (IEB) entails that in this case Ann must believe that she is scratching her neck. In other words, (IEB) would impose metaphysical demands on agents that make it so that there are situations where they cannot but act incorrectly.

This, I take it, is already an unwelcome result. But it is reflection on the properties that lead to this result that show the strength of the objection. For, we must imagine that Ann is in a case where in nearby situations she would not scratch her neck, as she intends.

⁵⁷ Here, I am building on Greco's (2015: 768) insight that anti-luminosity arguments can be used to call into question connections between a cognition and a belief by appeal. But whereas Greco suggests doing so by appeal to a Stalnakerian account of belief, I do so by appeal to a normative principle.

Presumably, however, this is something that she could be aware of, as Sisyphus could be aware that there was a point at which his mopping of the floor would not amount to a cleaning. Of course, she might not know—and possibly couldn't know—whether she is in such a case; but knowing about the possibility of so being is enough to make her realize that forming the relevant belief in such a case would be *risky*. Yet, are we to suppose that an agent can be so aware of the riskiness of forming a certain belief and yet not be capable of withholding belief? That would surely be surprising.

Of course, the objection is not decisive. Moreover, the objection doesn't generalize for a principle like (IARK). Then again, weakening the principles also weakens their explanatory power. In the previous chapter, I urged that we should not reject (PKP) just by appeal to counterexamples given the explanatory power of this principle (e.g. in explaining why an agent's saying (truthfully) 'I did not realize I was Φ ing' could be taken as grounds for concluding that the Φ ing wasn't done intentionally). Given that (IEB) and (IARK) lack the explanatory power of (PKP), I would now urge that objections that would be weaker against (PKP), such as objections by counterexample, should be taken more seriously as objections to (IEB) and (IARK).

The aim of this chapter is to present a general objection to cognitivism, one that applies to both (IEB) and (IARK), and should generalize for any form that a cognitivist principle might take. I begin by examining some reasons to hold a cognitivist view (Section 1). Next, I present my argument: both (IEB) and (IARK) suffer from counterexamples, and the counterexamples are shown to yield a surprising conclusion: cognitivist principles are tenable only if they are taken as evident analytic truths (Section 2). I close with some suggestions about how to move forward (Section 3).

1. ARGUMENTS FOR COGNITIVISM

Before presenting my objection to cognitivism, I would like to first give some reasons why one might be attracted to this view. After all, the claim that intentions are at least partly cognitive in nature is far from being a truth of common sense.

Many and varied reasons have been offered in favour of cognitivism, and I don't intend to consider them all. I shall focus my discussion on two related ones that have been offered in support of (IEB).⁵⁸ The first is that cognitivism offers the best explanation for the spontaneity of practical knowledge. The second reason is that cognitivism offers the best explanation for the fact that intentions are most naturally expressed by assertions.⁵⁹ I focus on these arguments because I think versions of them motivate a certain form of cognitivism, albeit not the standard one that has been defended in the literature. The presentation of them in this section is rather informal. More sophisticated versions of them will be presented in the next chapter as arguments against inferentialism.

With this in mind, imagine that you are eating breakfast with your flatmate on a Sunday morning, and she asks, 'Are you going to the gym today?'. Suppose further that, at the time at which she asks this question, you have not yet decided whether you will be going to the gym. You now consider the possibilities: if you do not go to the gym you can plan on having a full day of writing at home; on the other hand, if you go to the gym today, you wouldn't have to get up early tomorrow to do so. Let us say that on the basis of these considerations, you decide to go the gym, so you express your intention to your roommate:

⁵⁸ They would need to be modified to support (IARK).

⁵⁹ Versions of both arguments can be found in Velleman (1989) and Marušić and Schwenkler (2018).

‘I am going to the gym today’. Let us suppose that in this case you do in fact go to the gym, so that your assertion expresses knowledge.⁶⁰

This story, however familiar, can be used to present two powerful arguments for cognitivism. First, we are taking it that ‘I am going to the gym today’ expresses your intention. However, it is commonly taken that assertions express beliefs. So, the mental state expressed by ‘I am going to the gym’ is at least constituted by a belief. The cognitivist thus concludes that intentions that are expressed thus are at least constituted by beliefs. Moreover, since intentions are quite generally and naturally expressed as assertions, the cognitivist concludes that intentions just are, or are constituted by beliefs.

Second, consider how you would answer your flatmate’s question if instead of asking about what *you* would be doing, she had asked instead about what your brother was doing: ‘Is your brother going to the gym today?’ If that was the question, I take it that the natural way to answer it would be by looking for evidence about what your brother is likely to do (asking him, checking his calendar, etc.). For instance, if you know that he stayed up late the night before, and you know that when he stays up late, he tends not to go to the gym, that might be a reason to say that he is not, or not likely to go to the gym. Suppose the same is true of you: you stayed up late, and you tend not to go to the gym when you do so. In this case, it would be peculiar if you took these as the only grounds relevant for answering your flatmate’s question. But why? Given that your roommate asked about a

⁶⁰ Skeptics about knowledge of the future might balk at this claim. Note, however, that the example could equally well have involved an action that is immediately available to the agent (e.g. drinking the last bit of Orange juice). And it seems to me very implausible to deny that one can have knowledge of what’s going to happen in those cases.

matter of fact—what you are going to do—isn't it very reasonable to answer this question on the basis of evidence?

The reason we would find it peculiar if you were to figure out what you were going to do by looking at evidence in this way, is that what you do in such a case is *up to you*. As such, we expect you to figure out what you are going to do by making up your mind about it, where making up your mind consists in forming an intention. But the intention to go to the gym, as expressed in your assertion, amounts to knowledge in this case. On the assumption that knowledge is constituted by belief, this means that the intention must in turn be, or be constituted by a belief.

What shall we make of these arguments? There are problems with both, as stated. However, we shall be in a better position to recognize some of these problems once my objections to cognitivism are on the table. Hence, I shall delay discussion of them until the last section, and move on now to present my positive objection to the view.

2. COUNTEREXAMPLES GENERALIZED

Cognitivist principles seem to be subject to counterexamples. A famous one against (IEB) is the carbon-copier case from Davidson briefly mentioned in the last chapter:

[I]n writing heavily on this page I may be intending to produce ten legible carbon copies. I do not know, or believe with any confidence, that I am succeeding. But if I am producing ten legible carbon copies, *I am certainly doing it intentionally*. (2001b: 97)

In this example, a person has overwhelming reason to think that if she tried to make ten carbon copies, she would fail. Still, she gives it a shot and succeeds. As Davidson notes, the person in the example seems to both intend the action and act intentionally. However,

it seems equally plausible to think that all the while such a person failed to believe that she would be making 10 carbon copies while so doing, contra (IEB).⁶¹

It is partly on account of such counterexamples that Setiya defends as modest a version of a cognitivist principle as (IARK). As Sarah Paul has argued, however, even this modest principle seems to suffer from counterexamples like the following (I quote the complete example):

ANARCHIC HAND: Suppose an agent has a disorder something like Alien Hand Syndrome, which results in his hand acting “with a mind of its own.”⁶² Without the agent intending it, the hand frequently and capably carries out the action of buttoning up the agent’s shirt. Now, say the agent desires that his shirt be buttoned up and knows this will likely be done by his anarchical hand without him needing to intend it. However, his lack of control infuriates him, and he would greatly prefer that his hand button his shirt as a result of his intending it. But he knows from experience that intending to button his shirt with that hand actually lowers the probability that he will button his shirt, although it is still reasonably possible that he will. In light of all this, the agent’s confidence that he will button his shirt, or do anything toward this aim, should decrease as a result of his forming the intention. But if he is successful, he will have buttoned his shirt intentionally. (Paul 2009b: 553)

As Paul admits, the example is somewhat farfetched, but not enough to outstrip our concept of intentional action. Given this, take the case where, through his intention, the agent buttons up his shirt by moving his hands and fingers in the relevant ways. He would seem to be doing so intentionally, even though his doing so intentionally leads him to decrease (rather than increase) his confidence that he is actually doing so; and since he seems to do

⁶¹ One way to respond to this argument is to reject the intuition that the person in the example could be intending to make ten carbon copies while not believing that he will do so. However, studies conducted over Amazon Turk suggests that the intuition is widespread and robust, so I take it to be safe. The results of these studies are available upon request, and I hope to publish them in the future.

⁶² Paul seems to have in mind the phenomenon of Anarchic Hand Syndrome, rather than Alien Hand. Here’s Marcel on the difference: “In Hemisomatoagnosia or Alien Hand, one hand does not feel to be one’s own, especially when held with the other hand. This is a sensory phenomenon and has little to do with movement, whereas in Anarchic Hand it is certain actions performed by one hand that are disowned, not the hand itself” (2003: 76).

this as a basic action, without doing anything else as a means, this seems to be a counterexample to (AEC).

In his reply to Paul, Setiya observes that the case is a counterexample to (AEC) only under the assumption that the alternatives relevant to assessing the agent's confidence are ones where (i) the agent fails to form the relevant intentions, and (ii) his autonomic motor system is active (2009: 130-1).⁶³ However, he contends that the alternatives relative to which we should evaluate the agent are instead ones where (i) holds, but (iii) the person's autonomic motor system is inactive. On these grounds he concludes that ANARCHIC HAND is not a counterexample to (AEC).

However, this response does not address the heart of the matter, for we could stipulate that whenever the agent's autonomic system becomes inactive, some other condition ensures that there is an increase in the likelihood that the agent will button up his shirt. For instance, perhaps an evil scientist has inserted a mechanical device in the person's brain that activates when (i) and (iii) are met, a device that sends neurological signals through his body that reliably ensures that he buttons up his shirt when these conditions are met. If the agent knows this, it seems he could be more confident that he would button up his shirt and move his hands accordingly in a scenario where (i) and (iii) are met than when he intends to button up his shirt. But it would still be the case that if he succeeded in

⁶³ Setiya also calls into question whether the agent in the example actually fails to raise his confidence that he is acting in the relevant ways *while acting*, rather than just lose confidence ahead of time that he will act in the relevant way. But the example can be easily be modified to account for this. For instance, we could stipulate that the agent has been blinded and has an extremely underdeveloped sense of touch and proprioceptive system, so that he has no way of keeping track of his actions while he is acting.

buttoning up his shirt as a result of intending to do so he would be acting intentionally, contra (AEC), now understood with the relevant alternatives as Setiya posits.⁶⁴

There is a more general point here, which is that no matter which possibilities are taken as relevant to evaluating cognitivist principles, we can always come up with new counterexamples to them. For we can give a general recipe to find counterexamples of this type.⁶⁵

All traditional cognitivist principles have the following form (or entailment):

Cognitivist-Schema: (Necessarily) If a subject *S* has some practical attitude *I*, then *S* has a cognitive attitude *A* that is more cognitively committed to some proposition *p* than the attitude she would have if it was the case that *S* didn't have *I* and conditions *c*₁, *c*₂, . . . , *c*_{*n*} held.

Let β be a case that meets the following five conditions:

- I. In β , *S* has some arbitrary practical attitude *I*;
- II. The world is such that whenever *S* is in a case such as β , her having attitude *I* will decrease the chance of its being the case that *p*;⁶⁶
- III. The world is such that necessarily, if *S* lacks attitude *I* and *c*₁, *c*₂, . . . , *c*_{*n*} hold, the chances of *p*'s obtaining relative to any possibility where she has *I* will significantly increase;
- IV. *S* believes I, II, and III;
- V. *S* doesn't have access to any other relevant feature of the situation that bears on whether *p* will obtain.

Given these conditions, it seems that in β , *S* can (and perhaps should) decrease her cognitive commitment with regards to *p* relative to the circumstances where she lacks *I* and *c*₁, *c*₂, . . . , *c*_{*n*} hold. After all, she knows that the chances of *p*'s obtaining are higher in the latter than in the former case. Now, that conditions such as I and II can obtain simultaneously

⁶⁴ Again, the example is farfetched, but we are forced to consider farfetched examples when evaluating as weak a principle as (AEC).

⁶⁵ As it should be evident, the argument here is inspired by Zagzebski (1994), who argues that there is a general recipe for generating Gettier cases to call into question reductive analyses of knowledge.

⁶⁶ It does not matter for the purposes of the argument whether the probabilities are objective or subjective.

seems clear from a case like Paul's, and Setiya grants this; but it seems equally possible that 'philosophers have arranged' (to use Foot's (1967) famous phrase) so that condition III is also met. Thus, we may assume that some mechanism or other ensures the following: whenever *S* lacks *II* and c_1, c_2, \dots, c_n obtain, the chances of *p* increase relative to the case where *S* has *II*. Indeed, even if one disputes whether such conditions could obtain (because of some putative metaphysical truth), all the argument requires is that a person could believe that they obtain, even if they could not obtain.⁶⁷ Finally, IV and V should be uncontroversial. Since *II* is arbitrary, this means that it will always be possible to find a counterexample to cognitivism.

Here is another way to put the point. Take some feature of the world, *F*, that is taken as good evidence to think that *O* exists, since generally (or even necessarily) *F* results in some outcome where *O* exists.⁶⁸ It seems, given the nature of evidence, that there could be a case where there is enough counterevidence to make it possible (and rational) for an agent to believe that *O* is more likely to obtain when *F* is absent than when *F* is present. At the least, this seems possible whenever it is not analytic that *F* entails *O*. So, unless it is an analytic truth that *F* entails *O*, it is possible that there could be enough counterevidence that would lead one to think that it is more likely that *O* will not obtain than that it will obtain, given that *F* does.

The above is a general point about the nature of evidence. Let us now apply it to cognitivism, by considering the relationship between a practical attitude, *II*, and some proposition *p*. Unless it is analytic that *II* entails *p*, there will be cases where an agent would

⁶⁷ I owe this point to John Schwenkler.

⁶⁸ Here, 'evidence' is used broadly: you may substitute 'epistemic reason' if you prefer.

have more evidence to believe that p would obtain if she lacked II than if she had II . In those situations, an agent could (and should) presumably be more confident that p would obtain if she lacked II . What this shows is that counterexamples like the carbon copier and ANARCHIC HAND are symptomatic of a deeper problem, namely, that the only way to save traditional cognitivism is by claiming that cognitivist principles like (IEB) and (AEC) are analytic, which is surely implausible. Indeed, the cognitivist is committed to an even stronger and more implausible claim, namely, that the connection between practical and cognitive attitudes is clearly analytic.⁶⁹ After all, I can doubt even a complex mathematical proposition that is analytically true, because I can believe that there is good evidence against it. In light of these problems I think traditional cognitivism is untenable.

3. BEYOND STANDARD COGNITIVISM

A standard objection to cognitivism is to point to counterexamples. The standard strategy to deal with these objections is to weaken cognitivist principles to account for these objections. In the previous section, however, I suggested that this strategy is futile, since we have reason to think that *any* version of cognitivism will face such counterexamples, and that the only way to avoid them is to hold what is surely implausible, that cognitivist principles are clear analytic truths.

The problems raised by the counterexamples we have been considering go deeper. For they also call into question some of the central motivations for cognitivism.

⁶⁹ I owe this point to Josh Knobe.

Recall the two motivations for cognitivism mentioned in the first section: cognitivism is required to explain how an agent can spontaneously make up her mind about what she is going to do, and it is required to explain the assertoric nature of expression of intention. If these examples are to provide support for cognitivism (in its standard form), we must assume that they obtain with full generality. That is, we must assume that *whenever* an agent forms an intention spontaneously on the basis of deliberation, she makes up her mind in the sense of forming a belief about what she is going to do (or is doing); and we must assume that *whenever* an agent forms an intention, she is in a position to express it as an assertion.

However, the counterexamples considered also call into question the universality of these phenomena. Take the case of the carbon copier. The carbon copier spontaneously forms the intention to make 10 carbon copies; but he does not make up his mind to the effect that this is what he is going to do. By the same token, and given that this is his state of mind, it would be odd if the carbon copier expressed his intention by saying ‘I shall make 10 carbon copies’ or ‘I am making 10 carbon copies’. After all, he highly doubts that his action will take place (or is taking place). We would thus expect the carbon copier to express his intention rather by saying something like, ‘I intend to make 10 carbon copies’ or ‘I am intending to make 10 carbon copies’. The purpose of such weakened assertions would be precisely to avoid the suggestion that he knows or believes that he is making or will make 10 carbon copies.

This is not to say that the phenomena that motivate traditional cognitivism are not in need of explanation. In the next chapter I argue that to account for these phenomena we must hold that agential knowledge is non-inferential. And I shall go on to defend a form of

cognitivism in Chapter 5, although of a very different sort than the views considered in this and the previous chapter.

Chapter 3: Against Inferentialism

If there are no necessary connections between practical and cognitive entities, it is natural to think that agential knowledge is inferential. When I know, for instance, that I will go to the gym after lunch, or that I am typing right now, I know this because I have inferred this from other beliefs I have. There are two immediate *prima facie* problems with such an account, however. First, agential knowledge seems to be acquired *immediately* on the basis of making a decision, whereas on the inferentialist view it turns out that its acquisition is mediated by the beliefs that form the bases for the inference. Second, as Anscombe insisted (1958: 1-7), there seems to be an important distinction between predictive knowledge and agential knowledge. Thus, for instance, there seems to be an important difference from the way in which I might know that *you* are going to the gym later (on the basis of looking at your calendar), from the spontaneous way in which I know that *I* am going to the gym later (on the basis of having decided to do so). Yet, it is hard to see how an inferentialist account could draw a significant distinction between these two cases.

In this chapter I shall suggest that these problems are not merely *prima facie*. Although defenders of inferentialism have satisfactorily addressed simple versions of them, once they are properly developed, we shall see that the view lacks the resources to explain the characteristic directness and spontaneity of agential knowledge. Thus, I shall argue that we should reject inferentialism, and hold that agential knowledge is non-inferential.

I shall start by explaining in more detail what inferentialism is, focusing on what I consider the best version of the view, defended by Sarah Paul (2009a) (Section 1).⁷⁰ Next, I develop two objections to inferentialism: The first objection appeals to the assertoric expression of intentions to show that inferentialism fails to capture the characteristic directness of agential knowledge (Section 2). The second objection presents a dilemma that aims to show that inferentialism fails to capture the characteristic spontaneity of agential knowledge (Section 3). I close by generating a set of desiderata for an account of agential knowledge based on the results from the first three chapters.

1. INFERENCEALISM

There are many ways of developing the view that agential knowledge is inferential, depending both on one's view of what an inference is, and what particular kind of inference is involved in the acquisition of agential knowledge. To simplify the presentation of the arguments, I shall focus on the most sophisticated account in the literature, by Paul (2009a). However, I hope it will be clear by the end of the chapter that my criticisms of inferentialism generalize for any version of this view.

On Paul's account, agential knowledge is acquired in two stages. First, an agent takes advantage of her privileged access to her mental states to come to know that she has an intention to Φ .⁷¹ At the second stage, the agent makes an inference on the basis of this

⁷⁰ Other prominent inferentialists include Grice (1971), on whose work Paul is building, and O'Shaughnessy (1980). On the other hand, a view like Marušić and Schwenkler's (2018) (and likely Anscombe's), on which agential knowledge is gained on the basis of a *practical* inference is not called into question by my argument.

⁷¹ In her (2012), Paul presents an account of the mechanism by which we come to know our intentions on the basis of our decisions. This view might explain why knowledge of our intentions is distinctive and

knowledge. Paul argues that, generally and in the right circumstances, intending to Φ suffices for Φ ing.⁷² Similarly, for action in the future: in the right circumstances, an agent who forms an intention to Φ , will generally Φ in the future.

Given this general and reliable connection between intention and action, then, an agent can reason in accordance with the following schema:

Inferentialist Schema (IS)

- (1) I intend to Φ [in circumstances C_1, C_2, \dots, C_n].
- (2) If I intend to Φ [in circumstances C_1, C_2, \dots, C_n], then I am Φ ing/I will Φ .
- (3) Therefore, I am Φ ing/I will Φ .⁷³

A few remarks about (IS) are in order. First, (IS) represents a schema of the contents that certain beliefs must have to warrant drawing a certain conclusion. A belief with content (3) will be inferentially warranted by beliefs with content (1) and (2) only if the subject is also warranted in believing (1) and (2). However, to acquire such warrant further conditions are needed. For instance, the subject must *base* her belief in (3) on her beliefs in (1) and (2);⁷⁴ and her warrant for believing (1) and (2) should not be based on (3), on pain of circularity.

spontaneous. What I shall try to show, however, is that even so, the spontaneity does not carry over to knowledge of our actions when an inference is involved.

⁷² At least if Φ is of the right type (e.g. if there is a glass in front of you, and you form the intention to drink it, you will be engaged in the process of drinking it (note the progressive aspect)). As Paul notes (p.16) further complications arise when considering telic action descriptions, but I ignore these complications since they are irrelevant to the issue at hand.

⁷³ For simplicity, I shall take agential knowledge to be propositional, though see Thompson (2008: 120-27) and Hornsby (2016) for arguments to the contrary. The arguments in Stanley (2011, ch.3-5) for a propositional view of knowledge-how, however, seem to generalize for the case of intention, and seem to me to present a good case for a propositional view. However, this debate is orthogonal to the argument I shall give, so I shall not explore it further.

⁷⁴ This is a standard requirement on inferential warrant, though it is a difficult question what it takes for a belief to be appropriately based such as to meet that condition. See Boghossian (2014) and the reply by Wright (2014a) for two approaches to the issue, and discussions of the difficulties.

Second, the qualification about circumstances is in square brackets because there are more or less demanding ways of understanding this qualification. On a more demanding (internalist) account, the qualifications are part of the content of the beliefs that the agent employs in the inference. On a less demanding (externalist) account, it is sufficient if the relevant circumstances obtain at the time at which the agent draws the inference. Although Paul endorses the more demanding view, I want to remain neutral on this issue, since my argument is aimed at inferentialism generally, and it is consistent with inferentialism to endorse the less demanding view. Now, *which* circumstances are included in this qualification? Paul explicitly refrains from offering an exhaustive list (p.14), but she notes that plausible candidates include “knowledge of ability, conduciveness of circumstances, and one’s history as an agent [who reliably does what she intends]” (p.15), as well as knowledge that “many action descriptions function in such a way as to apply truly in large part because that description is what the agent intends” (p.16). Yet, inferentialist views might differ with respect to which circumstances they take as relevant, so I also want to remain neutral in this regard. Once again, though, I shall assume that the circumstances cannot include the conclusion the inference aims to provide warrant for. That would make the inference viciously circular.

Inferentialism economically explains agential knowledge, relying on the independently plausible claims that we have privileged access to our intentions, and that knowledge acquired on the basis of such access can serve as a basis for inferences. Yet, one might worry that inferentialism fails to capture the characteristic directness of agential knowledge since its acquisition is on this account inferentially mediated. Paul takes this objection head on, arguing that this directness is accounted for by the fact that inferential

reasoning “can take place rapidly and automatically at a non-conscious level, without the mindful entertaining of premises or feeling of drawing a conclusion” (p.10). One could object that agential knowledge should not be mediated even by unconscious processes, but absent further reasons to believe this, the objection is unpersuasive.⁷⁵

A different line of objection has been pursued by Setiya (2008, 2007). It begins with the contention that there are necessary connections between practical and cognitive entities, as captured by (IEB) and (IARK) in the previous chapter. If any such necessary connection holds, Setiya argues, inferentialism is false. The reason is that according to inferentialism the connection between practical and cognitive entities must be mediated by an inference, but since we can always fail to draw an inference, the connection between the two entities turns out to be contingent rather than necessary on an inferentialist view.

The problem with this argument is that, as we saw, there are good reasons to reject cognitivist principles of the sort appealed to in the argument. Indeed, Paul (2009b) herself rejects these principles, and she takes it rather as an *advantage* of inferentialism that it does not require such necessary connections (Paul 2009a). Hence, Setiya’s argument is at best dialectically weak.⁷⁶

Nevertheless, I shall argue that inferentialism should be rejected, and that practical knowledge is non-inferential. I shall not dispute that agents can gain knowledge of their actions on the basis of an inference. Rather, what I shall claim is that when they do so, this

⁷⁵ This reply can also serve as the basis to answer O’Brien’s (2007) objection that “it does not *seem* to me that I know I am raising my arm as a result of inferring that I am from my knowing that I intend to raise my arm” (p.177). The unconscious shape that inferences can take might make knowledge of our actions seem immediate without their actually being so.

⁷⁶ It is strong only against a view like Grice’s (1971), who upheld both inferentialism and cognitivist principles along the lines needed for Setiya’s argument.

knowledge is not direct nor spontaneous (in senses that will become clear shortly). Hence it is not agential knowledge.

2. CHATTY CHARLIE'S ASSERTIONS

We saw in the previous chapter that a common argument for cognitivism appeals to the assertoric form of expressions of intention. The cognitivist argues as follows:

1. Intentions are naturally expressed as flat assertions to the effect that one is doing or will do something.
2. Assertions express beliefs.
3. Therefore, intentions are (or are constituted by) beliefs that one is doing or will do something.

In the previous chapter I noted a problem with this argument. As stated, the argument is not clear about the generality of the first premise. It is ambiguous between a universal and a generic claim, to wit:⁷⁷

1*. It is *always* the case that: intentions are naturally expressed as assertions...

1**. It is *generally* the case that: intentions are naturally expressed as assertions...

The problem identified in the previous chapter is that the cognitivist needs 1*, but there are cases that suggest that only 1** is true. The cognitivist needs 1* because if there are cases where intentions are not expressible as assertions, that would suggest that the connection between intention and belief is merely contingent, rather than necessary. It would suggest

⁷⁷ The same is true of premise 2. Levy (2018) presents a powerful objection to the traditional argument against cognitivism by noting that assertions can express attitudes other than belief. The argument below aims to avoid some of the central problems Levy identifies in the traditional argument. As I note below, however, I disagree with Levy's claim that expressions of intention do not give rise to the same type of Moorean paradoxes that other assertions do.

that there are cases where an agent intends to do something but lacks the corresponding cognitive state pertaining to her action. As we saw, however, there seem to be cases where precisely this holds, and intentions are not naturally expressed as assertions.

There is another problem. Even if we granted I*, the argument would remain invalid. The reason is that a single speech act can express more than a single mental state. ‘I *know* that!’ can express both one’s belief that one knows as well as one’s exasperation; but we are not at all tempted to hold that exasperation is a species of belief. Similarly, it might be that the assertions by which we express intentions express beliefs without this entailing that intentions are (or are constituted by) beliefs.

Even so, I think that the fact that intentions are standardly expressed as assertions tells us something important about the nature of these mental states. To draw out the importance of this connection, however, we should focus on the normative profile of assertions. I shall argue that such a change of focus results in a powerful argument against inferentialism.

Assertion is widely held to have an epistemic norm: it is correct for a person to assert *p* only if she is in a certain epistemic state with respect to *p*. The most widely endorsed norm is the knowledge norm:

Knowledge Norm of Assertion (KNA): a speaker can legitimately assert *p* only if she knows *p*.⁷⁸

Because (KNA) will also play an important role in an argument below, let me briefly summarize the reasons that are taken to support it before presenting the argument. The first

⁷⁸ Defenders of KNA include Unger (1975), Slote (1979), Williamson (2000), DeRose (2002), Reynolds (2002), and Hawthorne (2004), among many others. Critics of the knowledge norm include Weiner (2005), Douven (2006), Lackey (2007), McKinnon (2013, 2015), Hill and Schechter (2007), and Schechter (2017).

reason to hold (KNA) is that this norm seems to best account for some of our conversational interactions. For instance, when someone asserts p , it is legitimate to ask, ‘How do you know p ?’, and if she says that she doesn’t know, it seems the speaker should recognize the impropriety of her assertion. These interactions suggest that as speakers we presuppose (KNA) and judge assertions by this principle. The second reason is that this norm seems to best account for certain Moorean paradoxes. Assertions of the form $\lceil p$ but I don’t know that p \rceil seem ‘clashing’ DeRose (2002), in some sense. For instance, take the sentence, ‘I am in New York, but I don’t know whether I am in New York’. The clash cannot be explained semantically, since the conjuncts are actually consistent, as is shown by the fact that I can perfectly well assert, of someone else, that she doesn’t know p even though p is the case. (KNA) explains this puzzling phenomenon. For according to (KNA), a person cannot legitimately assert the first conjunct without knowing that she is in New York. However, she cannot assert the second conjunct without knowing that she does not know that she is in New York; so, by the factivity of knowledge, she cannot legitimately assert this second conjunct unless she does not know she is in New York. One cannot both know and not know something, so Moorean assertions are necessarily illegitimate by the lights of (KNA). This explains the ‘clash’. These arguments seem to me to provide strong reasons to endorse (KNA).

With that in mind, consider the case of a man who cannot keep his thoughts in his head: whatever he thinks, he expresses out loud. Let’s call him ‘Chatty Charlie’. If you were to run into this peculiar being at the store, you would hear him voicing out a lot of information. For instance, you would hear him voicing out all his perceptual beliefs: ‘there are tomatoes over here; there are apples over there; the floor is white; this tomato feels soft;

it's so cold in here!' etc. In addition, of course, you would hear him voicing out a lot of information pertaining to his plans: 'I still need to find eggs; where are the eggs?; I'll walk over there; should I get one box or two?' etc. Like us, Chatty Charlie processes information like this at amazing speeds, so you might have to imagine that he speaks really fast, or perhaps that, unlike those of us who speak in a single voice, Charlie speaks in several voices at once, like a chorus (if you imagine him thus, you can also imagine that we have a machine that allows us to disentangle everything that Charlie says).

Whatever else we might need to imagine, the important point is that for Charlie there is no meaningful distinction between what he says and what he thinks. There is a single activity. But it bears emphasis that this is a *thinking* activity. Thus, like us, Chatty Charlie often makes inferences, both theoretical and practical. It's just that when he does so, he does so out loud.

Much of our thought is never communicated. And there are powerful Wittgenstenian arguments against the claim that it could all occur privately in this way. However, there is no good reason to suppose that it could all occur *publicly* as with Charlie. Hence, however remarkable, I take it that Charlie is a possible being, and one very similar to us. The point of focusing on his case is that it allows us to evaluate a person's thoughts in a direct and transparent manner, without worrying about the possibility that the correctness of our evaluations depend on features that stand outside our purview.

With this in mind, imagine now that Charlie was in front of you and you heard him reasoning as follows: "I am thirsty; here is a glass of water that would satisfy my thirst; I can drink this glass of water; I shall drink this glass of water" (he says, and he begins to

move his hand towards the glass, and then begins to chug the water). Suppose, further, that none of his other thoughts are relevant to his drinking of the water.

How should we evaluate the concluding assertion ('I shall drink this glass of water')? It seems clear that it is in good standing. After all, drinking the glass of water is under Charlie's control. If (KNA) is correct, therefore, we can infer from the legitimacy of this assertion that Charlie possesses knowledge. However, this intuitive verdict is in conflict with inferentialism. This is because Charlie has not gone through an inference of the sort that would enable him to know what he is going to do (we know this because, *ex hypothesi*, he would have expressed the inference if he had gone through it). Hence, Charlie does not know that he is going to drink the water, which means that it is incorrect for him to assert that he will. Because it predicts, wrongly, that such an assertion is incorrect, we should reject inferentialism.

One could object that the case is too farfetched to draw safe conclusions. However, the farfetchedness of the case does not, on its own, challenge the standing of the conclusions drawn from it. After all, the case is one where there is a clear normative judgment: *Charlie asserts correctly*. And clear normative judgments are not called into question by the farfetchedness of the case of application. For instance, consider the judgment that Voldemort acted wrongly when he used an *Avada Kedavra* spell to kill Harry Potter's parents. This judgment concerns a very farfetched situation, one where humans are able to kill one another through the use of magic (unlike Chatty's case, I take this to be impossible given the laws of nature). However, the fact that such a possibility is so farfetched does nothing to call the normative judgment into question. Like the judgment

that Voldemort acted incorrectly in killing Harry's parents, our judgment that Charlie asserts correctly seems to me perfectly safe.

Another objection is that the concluding speech-act is not a genuine assertion. This would mean that (KNA) is not applicable to it. But what does that mean? The speech-act surely *sounds* like an assertion; and it behaves exactly like one. For instance, it can be the proper origin of a testimonial chain, which seems to be the central function of genuine assertions. For instance, suppose you had turned away from Charlie, but heard him saying that he will drink a glass of water. You could then tell someone else that that's what Charlie was doing on the basis of that assertion. You could do so because you would know this; and you would know this, presumably, on the basis of Charlie's knowing it as well. Moreover, if (KNA) didn't hold with respect to it, we would expect that Moorean paradoxes wouldn't arise for it; but they clearly do. If Charlie had asserted instead: 'I shall drink this glass of water, but I don't know whether I shall drink this glass of water', this would 'clash', showing that we apply KNA to it just as we do to any other assertion.⁷⁹ There is therefore no reason to think that this is somehow a different kind of speech-act than it appears to be at face value.

Charlie's case forces a hard choice on the defender of the indirect model: either reject the knowledge norm of assertion or accept the direct model. If she takes the first option, the defender of the indirect account incurs a theoretical debt, owing us an alternative explanation of the data that supports the knowledge norm. Because I doubt that an acceptable alternative is viable, I think we should reject the indirect model.

⁷⁹ My judgments here are in opposition to Levy (2018).

3. ALIENATED KNOWLEDGE

The second objection against inferentialism appeals to what I take to be an essential property of agential knowledge that I shall call ‘spontaneity’. Agential knowledge is available from the practical point of view from which the action is regarded by the agent as *up to her*.⁸⁰ By contrast, ‘alienated’ knowledge, knowledge had from a standpoint from which the action is regarded a settled matter, is not agential knowledge. The addict’s reply, “I am going to smoke a cigarette today, I just know it—you see, I am addicted” is thus not an example of agential knowledge. An indication of this is that such a statement has the character of a mere prediction, rather than an expression of intention.

The objection is in the form of a dilemma. I argue that an inferentialist account either cannot explain how we ever know about our actions in a spontaneous way, or else it cannot explain how we ever attain knowledge of our actions rather than just our intentions. In presenting the dilemma, I begin with knowledge of future action, where the problems are most salient. This would be enough to call inferentialism into question, which purports to offer an account of agential knowledge in general, both knowledge of action in progress and knowledge of future action. As I shall argue, however, once we see that inferentialism has trouble accounting for knowledge of future action, it will be plain that it also has trouble accounting for knowledge of action in progress.

⁸⁰ This is connected to what psychologists nowadays call the ‘sense of agency’ and ‘sense of control’. See Pacherie (2007) for a discussion of these and related notions, and Marcel (2003) for important links with philosophical work on agential knowledge.

Consider, then, an ordinary case of knowledge of future action, such as my knowledge that I will take a walk after lunch. According to inferentialism, I know this on the basis of an inference with the following structure:⁸¹

(4) I intend to take a walk after lunch.

(5) If I intend to take a walk after lunch, I will take a walk after lunch.

(6) So, I will take a walk after lunch.

As noted above, one of the seeming advantages of inferentialism is its ability to explain agential knowledge by appeal to independently plausible claims. For, according to inferentialism, the knowledge that (4) and (5) give me of (6) is ordinary inferential knowledge, on a par with my knowledge of (9) on the basis of (7) and (8).

(7) The fire is burning in the chimney.

(8) If the fire is burning in the chimney, the room will warm up.

(9) So, the room will warm up.

Yet, there is a crucial disanalogy between the two inferences, pertaining to their conditional premises: whereas the connection in (8) between the fire's burning and the room's warming up obtains through an independent process that can run its course without any contribution on the part of the agent, the connection in (5) between the intention and the action is one that necessarily requires the person to execute it.⁸² Another way to put this is that the truth of the conclusions are oppositely related to the agent's involvement: (9) will obtain unless

⁸¹ For simplicity, I omit the qualification about circumstances. Let's assume they are met throughout.

⁸² The process must be the particular process of *execution* since otherwise the resulting action won't be intentional, even when caused by an intention with the appropriate content (Davidson 2001b; Frankfurt 1978).

the agent intervenes, whereas (6) won't obtain unless she does. This crucial difference gives rise to a dilemma.

First Horn

On the one hand, the agent might hold the belief in (5) on the basis of a mere empirical generalization, as one might hold (8).⁸³ She realizes that in situations such as the one she finds herself in, where, for instance, she knows she has the ability and opportunity to execute her intention, she generally executes it (just as, generally, when the fire is on, the room warms up). However, in an ordinary case it is up to the agent whether she executes her intention or not. Unlike the case of the fire, where nature simply needs to run its course, the action will not execute itself, so whether the conditional premise holds true depends on what she does.⁸⁴ True, not going for a walk (say, out of laziness) might constitute a form of practical irrationality, but the point is that it is up to the agent—the very one considering what she is going to do—to act rationally or irrationally, as she should recognize. To recognize this, to take the execution of intention as being up to oneself, is constitutive of the agential standpoint; to think otherwise, to take the execution as something that is simply settled given one's circumstances (even circumstances that include one's intentions), is a paradigm example of bad faith, taking an alienated standpoint on a phenomenon that is

⁸³ This seems to be the best way to ensure that the belief in the conditional claim is not dependent on belief in execution. The two horns of the dilemma at a more general level are as follows: either (a) belief in the conditional claim is not dependent on belief in the execution or (b) it is so dependent. (a) leads to alienated knowledge, and (b) to a regress (assuming inferentialism).

⁸⁴ This point bears on an objection that Sarah Paul raised: Can't the agent simply have it as a background assumption that she won't change her mind (as one might have it as a background assumption that there will be sufficient oxygen for the fire to keep burning)? I answer that she cannot: the agent would at least have to assume the stronger claim that her mind is set on doing what she intends, i.e. that she will execute her action (given the considerations in n14). More importantly, the same worry would arise for such a background assumption: on what basis would the agent assume (or be warranted in assuming) that she won't change her mind? Given that she is capable of changing it, she cannot from the agential standpoint assume that her intention settles that she will not change her mind.

within one's control.⁸⁵ Such an alienated standpoint can of course yield knowledge: the addict can, after all, come to know that she will be taking a hit before the end of the day by reflecting on her addicted urge. However, such knowledge would not be agential knowledge.

It might be objected that the addict's situation is crucially different from the situation of the person inferring that she will take a walk. After all, the addict's knowledge is not grounded in an intention. She has not yet formed an intention, and if she had, it would be unclear whether her knowledge is alienated in a way incompatible with agential knowledge. This difference is crucial, since intentions are plausibly taken as the loci of agency. Hence, one might argue that if an agent knows about her action based on knowledge of her intention (and knowledge that the relevant circumstances obtain) that suffices to guarantee that her knowledge of the action is spontaneous.

To see why those conditions are insufficient, consider the following case.

Compelling Intentions: Moria often forgets her plans, so she carries around an agenda where she writes down what she has decided to do as soon as she forms an intention. On any given day, you can read phrases like: *At 5pm I shall go to the bank*; or *After lunch, I shall go to the gym*. These notes have for Moria the force of absolute and irrevocable commands: provided that on the given day she has the ability and opportunity to execute these actions she will do so. Thus, if Moria read on a given day, *Today I will go to the gym*, she would not deviate from that plan even if you offered her a million dollars to do so. Once written down, the matter is settled for her. This is why these notes are such a reliable source of information. If you asked Moria what she was doing on a given day, she would check her notes and reason as follows: 'I wrote that I am going to Φ ; I generally do what I wrote down (when I'm able to, etc.); hence I am going to Φ '. On that basis, for instance, she might conclude and come to know that she is going to the gym later on.

It seems clear that Moria can generally acquire knowledge of her actions on the basis of what she reads in her diary. But it is also clear that this knowledge is not agential

⁸⁵ My understanding of bad faith is indebted to Moran's (2001) now classic work on the topic.

knowledge, since it is alienated. The knowledge is alienated despite being grounded in an intention that the agent forms while correctly assuming that the appropriate circumstances will obtain and that she will not change her mind about what to do. This shows that even knowledge of an action that is grounded on an intention can be alienated.

Moria's case shows that knowledge gained on the basis of an intention does not guarantee spontaneity. Granted, the case is far from standard. However, it is hard to see how the differences between it and standard cases could be exploited by an inferentialist theory to explain why Moria lacks spontaneous knowledge of her actions. In particular, notice that we can assume that Moria correctly writes what her intentions are because she has the special access to them that agents normally have. In addition, we could modify the case so that only a few milliseconds go by between the time at which Moria writes down her intention and the time when she reads her agenda. Moria's knowledge would be equally alienated in such a case. The central difference between Moria's case and an ordinary one is therefore in the manner in which the agents keep track of their intentions: agents standardly do so through memory, while Moria does it through her agenda. Clearly, though, this feature does not explain why Moria's knowledge is alienated. After all, I can have agential knowledge of my actions even if I have to check my calendar to remind myself what I have planned!

The intuitive reason why Moria's case differs from the standard one (including the standard one where we have to aid our memory by the use of an agenda) is this: when Moria reflects about what she is going to do, she does so from a standpoint that fails to recognize that what she is going to do depends on what she thinks about the matter *from that very standpoint*. Moria thus regards her actions from the standpoint of a mere

spectator, simply waiting to see how a certain process unfolds, albeit one that begins with an intention formed by her past self. But that is not the standard way in which we regard our intentions. Our intentions are standardly maintained through our ongoing commitment to act in a certain way, and the agential standpoint is the standpoint from which our actions are regarded as so dependent on that commitment. It is thus constitutive of the agential standpoint from which practical knowledge is available to regard the execution of an action as in our power.

The same is true of our actions in progress, which are sustained by our ongoing commitment to them. For this reason, these considerations are not restricted to knowledge of action in the future. An agent can regard what she is doing from the standpoint of a mere spectator, as a process that merely unfolds before her eyes. Alternatively, she can regard her ongoing action as one that is itself sustained by and dependent on her commitment to execute it. And her knowledge will be spontaneous only if she regards her action from this second standpoint.

The problem with taking (5) as based on a mere empirical generalization—whether to gain knowledge of future or ongoing action—is now clear. It requires the agent to regard her action from an alienated standpoint, since it requires her to think of the execution of her action as something inevitable given her previous intention. Since belief from this alienated standpoint would, on the inferentialist view, necessarily be part of the justification that one has for believing that one will act in a certain way, this means that on the inferentialist view one could know about one's actions only from an alienated standpoint. But the phenomenon of agential knowledge is supposed to be special precisely in being knowledge from the agential standpoint. Hence, agential knowledge cannot be

gained on the basis of an inference relying on (5), if (5) is understood to hold just on the basis of an empirical generalization.

Second Horn

Suppose the agent's belief in (5) is held from a non-alienated standpoint, compatible with the recognition that the connection between her intention and her action is dependent on her actually executing her intention. In that case, it would seem that my knowledge of (5) would depend on my further knowledge that:

(10) I will execute my intention to take a walk after lunch.

Now, either one's execution of an intention to Φ is identical with one's Φ ing, or it is not. Suppose it is identical (as I think we should hold for reasons considered below). In that case my executing my intention to take a walk after lunch is identical to my taking a walk after lunch. But this would mean that (10)=(6). However, since on this view (10) is supposed to be the basis for holding (5), this means that the inference (4)-(6) is viciously circular, since the conclusion, (6), is based on a set of premises that includes itself. This shows that if execution and action are identical, and if one can have knowledge of conditionals linking one's intentions to one's actions from the agential standpoint, that knowledge cannot be prior to the knowledge of one's actions. This, of course, is incompatible with inferentialism.

Suppose, on the other hand, that the execution is not identical to the action. Then, in the case at hand we have a second action, viz. my execution of the intention to take a walk after lunch. However, how do I know that I will carry out *this* action; that is, on what basis do I hold (10)? There are two options: either the action is intentional or it is not. It would be hard to argue that it is not intentional: on the view under consideration the

execution stands in the instrumental relation characteristic of intentional actions (presumably, I execute my intention to take a walk in order to take a walk after lunch), and the action is under my agential control: if I stop intending to execute, I will stop executing my action. More importantly, holding that it is not an intentional action would not help the inferentialist, since taking this view would push us back into problems related to the first horn of the dilemma: the agent would have to regard her action as settled by what's not under her control (in the way characteristic of intentional action), namely, her execution, thus understood.

This only leaves one option open for the inferentialist: the execution of an intention to Φ is a different intentional action from the action of Φ ing. Since it is another intentional action, the inferentialist would have to hold that I know (7) on the basis of two further beliefs, namely:

- (11) I intend to execute my intention to take a walk after lunch; and
- (12) If (8), then (7).

Yet, the problems with (5) clearly re-emerge with regards to (9): how do I know this conditional claim? We can rule out that I know it as a simple generalization, since this would have me knowing (9) from an alienated standpoint. Hence, I must know it on the basis of the claim that I will execute this further action, i.e. on the basis of the further belief that:

- (13) I will execute my execution of my intention to take a walk after lunch.

Clearly, though, this gives rise to the same problems as before, and we are caught in a vicious regress, since we are unable to give epistemic grounds that would provide warrant for (5).

One could reply that this is a problem for everyone. For the problem arises once we realize that to culminate in an action, the intention needs to be executed. If we take execution as an intentional action (as we seemingly should), then the agent must intend to do it as well; but then she must execute this further intention; and so on. However, this is only a problem if we think that the execution of the intention is not the same as the action. Avoiding such a regress seems to me a good reason to hold the identity claim instead. Yet, as we saw in the discussion of the second horn, the inferentialist cannot endorse this identity without making the inference circular; but this is only because the inferentialist holds that knowledge of a conditional claim like (5) is more basic than knowledge that one will act. I suggest we reject that assumption.⁸⁶

Of course, the inferentialist could bite one of the bullets. For instance, she could hold that agential knowledge is available only from an alienated standpoint. This, to my mind, would be to recognize that the phenomenon that the inferentialist identifies is not agential knowledge. Another option would be to admit that there is a regress but deny its viciousness. This would, at a minimum, require an endorsement of epistemic infinitism, the position that a belief might be warranted on the basis of an infinite inferential chain.⁸⁷ Infinitism is an unpopular position, and commitment to it seems to me too steep a price to

⁸⁶ Thus, for instance, we could endorse a view like Setiya's (2011), on which knowledge of our intentions is posterior to knowledge of our actions; but the argument is consistent with a view that gives no priority to either knowledge.

⁸⁷ I say "at a minimum" because even infinitists recognize forms of epistemic regress that are vicious. Hence, the inferentialist would also have to show that the present regress is among the non-vicious ones.

pay.⁸⁸ The inferentialist's options for responding to the dilemma thus look grim. I therefore conclude that agential knowledge is non-inferential.

4. TAKING STOCK

This is a good place to summarize the results of the first three chapters. In the first two chapters I presented arguments against cognitivist principles to the effect that there are necessary connections between practical and cognitive entities. In this chapter, I argued against an alternative view, inferentialism, according to which agential knowledge is inferentially acquired. I suggested that this view cannot explain the characteristic (epistemic) directness and spontaneity of agential knowledge. Although the central conclusions of these chapters are thus negative, there are also some positive results. In particular, the arguments thus far suggest a number of desiderata for a successful account of agential knowledge. They are as follows:

1. *Important connection between intentional action and practical knowledge:* As I emphasized at the end of Chapter 1, the fact that it is possible to act intentionally without knowing what one is doing is consistent with the view that there is an important connection between intentional action and agential knowledge. An account of agential knowledge should (a) explain what this connection consists in, and (b) explain what it is about intentional action and agential knowledge that makes it so that such a connection holds.

⁸⁸ Peter Klein has been the foremost defender of the view among contemporary epistemologists. See Klein (2011) for a helpful survey of the literature.

2. *Assertoric Expressions of Intentions*: It is not always correct or natural to express an intention by an assertion. However, the fact that intentions are *generally* best expressed assertorically—expressing the agent’s agential knowledge—is in need of explanation. A theory of agential knowledge should explain why intentions can, in the relevant situations, be naturally and correctly expressed by assertions.
3. *Directness*: An agent who deliberates about what to do and concludes that deliberation by making a decision seems to be in a position to know what she is doing (or will do), where this knowledge is epistemically unmediated. Even if agents are not always in a position to acquire knowledge from deliberation in this way, the case where they are so positioned raises a puzzle: What would enable an agent to acquire knowledge directly just by making a decision?
4. *Spontaneity*: Agential knowledge is practical, and practical knowledge is knowledge had from the standpoint where an agent regards what she does as up to her. That is, she must regard her actions as dependent on what she thinks her actions are (or will be). What kind of thinking could this be, though, such that it could at once be responsible for what happens in the world (in the way we tend to associate with desires and intentions) while at the same time amount to knowledge? This question seems particularly pressing if we reject the cognitivist view that what it is to be an intention is to be a belief, as I argued we should.

I shall aim to provide an account of agential knowledge that answers the questions and meet the desiderata in 1-4. I begin by pursuing the strand of thought suggested at the end of Chapter 1. The idea there was to think that we can account for the connection between intentional action and practical knowledge (point 1 above) if we model it on the connection

between perception and perceptual knowledge. Just as we can gain direct knowledge of our surrounding through our perceptions—so the suggestion goes—we can gain direct knowledge of our actions through our intentions.

I pursue this idea in the next chapter, by providing an account of how perceptions give us such direct knowledge. I shall then use this account in Chapter 5 to argue that intentions can play an analogous epistemic role. As we shall see in that chapter, this will be only part of the story that needs to be given to account for all the features of agential knowledge. In addition, we shall need to endorse a certain kind of cognitivism, albeit one quite different from the traditional ones found in the literature.

Chapter 4: Knowledge through Virtuous Performances

I have argued that both cognitivism and inferentialism fail to provide us with a plausible account of agential knowledge. In the next chapter I shall develop an alternative account that aims to meet the desiderata set out at the end of the previous chapter while avoiding the problems with the other two views. One of my central claims is that beliefs about our actions are directly warranted: their warrant does not depend on other beliefs. This chapter lays the groundwork for the first thesis by presenting a general theory of direct warrant.

The chapter is divided in two sections. First, I present a brief, traditional explanation of why we need a theory of direct warrant as the only plausible way to hold a foundationalist account of knowledge, and explain why classical reliabilism is not up to the task (Section 1). Second, I defend a virtue-theoretic account as the best supplementation to such a reliabilism, and use it to explain how we acquire perceptual knowledge directly (Section 2).

Before we start, however, it will be helpful to explain the terminology I shall be using. In what follows, I shall speak in terms of epistemic ‘warrant’. A belief is warranted when it is correctly formed and maintained.⁸⁹ Warrant is thus a normative property that must be met for a belief to be good epistemic standing, and thus to amount to knowledge. Some may prefer to speak of ‘justification’ here; but, in a similar fashion as Burge (1993, 1997, 2003, 2010), I reserve that term for the particular kind of warrant that obtains when

⁸⁹ I thus take warrant to be primarily a property of the process of holding beliefs, and only derivatively of the belief state itself. With that in mind, there is no harm in speaking of beliefs themselves as warranted as I shall sometimes do in what follows.

someone has access to the grounds that ensure that the belief is in good epistemic standing.⁹⁰ The notion of warrant is meant to be broader, since a belief can be warranted given certain features of its source that are not accessible to the person. Justification is thus only a species of warrant (an internalist one).

It is also important that in my view, although warrant is necessary for knowledge, further conditions (in addition to factivity) may need to be met.⁹¹ For instance, some may be inclined to think that anti-defeat conditions must also be met: even if a belief is warranted, it might not constitute knowledge, because the person has good reason to think that it is false or unwarranted (when it isn't in fact so).⁹² Additionally, I take it that you might form a belief on the basis of the right source (so that it is warranted), and still fail to obtain knowledge because the environmental conditions are epistemically hazardous. The man who sees a real barn in the midst of a field full of fake ones is warranted in having the belief that in front of him there is a barn; but he does not know this fact. However, nothing I say hangs on this, and, for simplicity, I shall focus on cases where there are no defeating nor similar undermining epistemic conditions.

⁹⁰ More carefully: Burge uses 'justified' in the restricted way in which I use 'justification', but he sometimes uses 'justification' more broadly to encompass even beliefs that are not justified (such as the beliefs of animals). I follow Graham (2012, 2014) in keeping these terms more closely connected.

⁹¹ Since I am inclined to agree with Williamson (2000) that knowledge is basic, I doubt that we can give a set of reductive conditions that would be jointly sufficient for knowledge. I am here departing from Plantinga's (1993a) usage: on Plantinga's view, warrant is the ingredient that turns true belief into knowledge.

⁹² However, see Lasonen-Aarnio (2010) for a strong argument against this view. In Chapter 7 I side with her view.

1. FOUNDATIONALISM AND THE NEED FOR A BETTER RELIABILISM

To constitute knowledge, a belief must be warranted. Many beliefs are warranted on the basis of other beliefs. For instance, my belief that there is a beer in my fridge is warranted on the basis of the further beliefs that I put a beer in my fridge last weekend, and that no one has drunk it. Yet, not all beliefs can be warranted in this way. For suppose every belief were warranted on the basis of further beliefs (the ‘warranting beliefs’). The set of warranting beliefs would either include or not include some of the beliefs they are supposed to warrant. If the warranting beliefs include some of the warranted beliefs, then there is circularity in the structure; but circularity is incompatible with warrant. And if the warranting beliefs do not include any warranting beliefs the chain of warrant would have to be infinite; but infinite chains of beliefs do not provide warrant.

Of course, the above are two of the horns of Agrippa’s famous skeptical trilemma. Some seek to disarm the trilemma by biting the bullet on one of these two options. Thus, coherentists argue that at least in some cases circularity is compatible with warrant (“if the circle is big enough”); and infinitists contend that at least some infinite chains are compatible with warrant. This is not the place to argue against these views. I do not find either to be plausible, so I shall set them aside.

The third horn in the traditional trilemma is that our beliefs are unwarranted, since they are not based on any further belief. For, it is assumed, if a belief is not based on some further belief, then it is baseless, and baseless beliefs cannot be warranted. Yet, this assumption is questionable: even if many of our beliefs are warranted insofar as they are based on further beliefs, some beliefs might be warranted in a different way. The foundationalist contends that there are some beliefs that are warranted without being based

on further beliefs. Such beliefs provide the bases for forming further beliefs. They are foundational.

A central challenge for foundationalism is therefore to provide an account of such foundational beliefs that answers the following two questions: First, what beliefs are foundational in this way? Second, in virtue of what are such beliefs foundational? The empiricist and the reliabilist traditions in epistemology provide two classic answers to these questions. According to empiricism, perceptual beliefs—beliefs formed on the basis of our perceptual systems—are the central bases of our knowledge.⁹³ And according to reliabilism, a belief can be foundational just in case it is formed on the basis of a process that reliably yields true beliefs.

Empiricism, as I have formulated it, is now thought to be excessively narrow. Most would agree that in addition to perceptual beliefs, we should include core logical and mathematical beliefs. And many today would agree with Reid that we should include introspective beliefs, recollective beliefs, and testimonial beliefs among those that can be foundational. My sympathies lie with this pluralist, Reidean tradition. However, in this chapter I shall take perceptual beliefs as the paradigm instances of foundational beliefs, and thus as the central test case for an account that attempts to answer the second question about the features that make them so.

A central advantage of reliabilism is that it explains why perceptual beliefs are foundational. Our perceptual systems make information about our environment available

⁹³ Classical empiricists would identify perceptual experiences, rather than perceptual beliefs, as the central bases of our knowledge, where experience is understood as a luminous mental state with internally-defined content. This view has fallen into disrepute in the latter half of the past century. This, in my view, is for good reasons that I am not here able to consider.

to us, information that is reliably true. Hence, forming beliefs on the basis of our perceptual systems is forming beliefs that are reliably true. It follows from reliabilism that perceptual beliefs are warranted. Yet they are not warranted on the basis of other beliefs. In the right circumstances these perceptual beliefs constitute knowledge, thus constituting a proper foundation for our system of beliefs.

However, classical reliabilism faces a serious problem, stemming from the fact that reliability is too weak a notion. Thus, reliabilism is incapable of explaining why certain beliefs that seem unwarranted are so, yielding false positives. To see this, consider the following case from Plantinga:⁹⁴

BRAIN LESION: Braid has suffered a brain lesion that causes him to believe that he has a brain lesion. This means that Braid's belief will be formed on the basis of a (perfectly) reliable source, since in nearby worlds where he forms that belief on the basis of this source (his brain lesion), the belief (that he has a brain lesion) is true. (Plantinga 1993b: 195-8, 205-07)

Because Braid forms beliefs on the basis of a perfectly reliable process, it seems that according to standard reliabilism these beliefs constitute knowledge. But this seems like the wrong verdict. It seems as though Braid forms beliefs that are still, in an important sense, only accidentally true in a way incompatible with knowledge. What the case suggests, then, is that reliability is too weak a notion to capture the non-accidental connection between truth and knowledge.

⁹⁴ Another famous case that is often used to support this point is Bonjour's (1985) case of Norman, a man who reliably forms beliefs about the president's whereabouts, without any knowledge that he does so or how he does so. The account presented below is able to explain why in *some* versions of the case Norman doesn't have knowledge. However, for the reasons presented by Srinivasan (ms), I rather take it as an advantage of the view that it doesn't rule out that in other versions of the case, Norman does have knowledge.

The counterexamples that affect reliabilism call into question the sufficiency of reliability to account for warrant, leaving necessity intact.⁹⁵ As such, I shall assume that what we need to do to avoid the counterexamples is to finesse the account, rather than look for something entirely different. There are a number of proposals in the literature that aim to do just this. In the next section, I shall defend a particular version of virtue reliabilism. Before doing so, however, it is worth noting that the virtue-theoretic account defended here is stronger than many forms of reliabilism in the literature, and it shares many features with other outstanding accounts of epistemic warrant such as proper-functionalism.⁹⁶ Thus, although in the next three chapters I appeal to the epistemic framework developed below, it would be possible to achieve similar results by appeal to other frameworks that share many similarities with my own.

2. WARRANT BY COMPETENCE

The account I shall defend is virtue-theoretic in nature in two related respects: (1) it takes epistemic performances as part of the broader class of telic performances, and (2) it explains epistemic warrant by appeal to the notion of competence.⁹⁷ Much of this section will be devoted to developing (2), but first I need to explain (1).

⁹⁵ Of course, there are objections to reliabilism, like the new evil demon problem (Cohen 1984; Lehrer and Cohen 1983), that purport to call the necessity into question. However, I think there are good responses available to the reliabilist to deal with this problem, along the lines of Littlejohn (2015) and Williamson (2015), though see Greco (forthcoming) for criticisms of this strategies.

⁹⁶ See e.g. Burge (1993, 2003); Burge and Peacocke (1996), Graham (2012, 2014), Greco (2003), Peacocke (2003), Plantinga (1993a), and Sosa (2010, 2007, 2015) for accounts of warrant with affinities to the one presented here. Some similarities and differences between these accounts and mine will be noted in the literature, but I shall not aim to provide a detailed comparison.

⁹⁷ These two commitments are central to the virtue epistemology of Ernest Sosa, and his influence on the account presented in this section should be evident to anyone acquainted with his writings (especially (2010, 2007, 2015)). There are, however, a number of important differences between our accounts. First, I aim to

Telific performances are those that by their very nature aim at a certain goal that determines standards of correctness for those performances.⁹⁸ They include sub-personal-level performances, like the operations of the cardio-vascular system which aims at properly regulating blood flow throughout the body, and personal-level performances, like the shooting of a golf ball aimed at sinking it into a hole. Epistemic performances are processes by which one makes up one's mind as to how the world is. They are partly defined by their distinctively epistemic aim.⁹⁹ For simplicity, I shall assume this aim is truth.¹⁰⁰ Because they have an aim, epistemic performances can be assessed for competence: they will be competent insofar as they achieve that aim (or tend to do so) well, in a sense further explained below.

An important feature of performances is their 'stackability': performances can be part of further performances, and contribute to the broader performance by playing a distinctive role within it. This point is familiar from discussion of actions: I crack the eggs as part of making an omelette, a performance which is in turn part of making breakfast. However, the point generalizes in two important ways. First, it also holds true of joint

give a unified and externalist account of knowledge, whereas for Sosa, knowledge bifurcates into animal knowledge and knowledge full well (which requires the right internal justificatory conditions). Second, I define competence in terms of guidance rather than taking it as a basic notion. Third, and finally, as I explain below, this definition allows us to remain neutral on debates on the nature of know-how.

⁹⁸ I use 'telific' instead of 'telic', which is already used to classify a certain class of verbal phrases in terms of their aspect.

⁹⁹ See Shah (2003) and Shah and Velleman (2005) for arguments to the effect that the formation of belief has this aim.

¹⁰⁰ Nothing hangs on this. Those partial to the knowledge-first program in epistemology (as I am) may wish to substitute knowledge instead, and the account below would work equally well. Miracchi (2015) defends a view along these lines, taking as basic our capacities to know (rather than the capacities to form true beliefs). An alternative along the same lines is to follow Williamson (2000) in thinking that the alternatives relevant for assessing for knowledge cannot be defined independently of the notion of knowledge.

actions: several workers can be engaged in different performances (painting, laying bricks, cutting wood, etc.), all of which together amount to the building of a house. Second, the point also holds true of some biological performances that aren't actions. Take the process of digestion, which includes as constituents the contractions of the esophagus by which the food is delivered to the stomach, where the food is broken down by enzymes to be delivered to the small intestine, etc. Here, again, we see a set of performances that are dependent on one another and add up to a larger performance: digesting.

With these points in mind, we can now return to the question of epistemic warrant. As noted, I am taking as a working hypothesis the virtue-theoretic view that epistemic warrant is definable in terms of competence: one is warranted in having a belief just in case the belief is competently formed (and maintained) by the standards appropriate to it given its alethic goal. The question, then, is: What does it take for a belief to be competently held in this way? An advantage of formulating the question in these terms is that we can recognize it as a specification of a more general and more tractable question, namely: What does it take for a telic performance in general to be competently done?

To answer this question and to see how competence differs from mere reliability, consider a waiter working at a restaurant where all the wines in the menu pair well with every food item in it. Whenever customers ask which wine they should pair with their food, the server reliably makes a good recommendation. However, the server knows nothing about wines, and essentially chooses at random. Hence, although this server reliably makes good pairing recommendations, it is clear that he doesn't do so competently: there is an important sense in which this achievement is accidental, primarily explained by features external to his performance.

The waiter's recommendations do not seem to be geared towards making correct recommendations. This is what seems to explain the lack of competence. Consider, on the other hand, a waiter who either through knowledge or natural ability reliably recommends a correct pairing. Such a waiter would seem to make the recommendations competently, because his actions are guided by the aims supplied by her knowledge or natural ability.

What we need, therefore, is an account of the kind of guidance that knowledge and intuition exercise in such cases. I suggested the following:

Guidance: A performance p is guided by some aim A iff (a) the correctness of p is determined according to the extent to which it tends to achieve A , and (b) p is structured to adjust so as to achieve A .¹⁰¹

In turn, I suggest defining competence as follows:

Competence: A performance of p S is competently done relative to some aim A iff (i) p is such that it reliably achieves A , and (ii) the reliability of the performance is explained by the fact that p is guided by A .

The initial waiter's recommendations in the example meets condition (i); but he doesn't meet condition (ii) since he fails to meet (b) in the guidance condition. His performance is not structured to adjust so as to accomplish the aim of making good recommendations. If the menu were to be changed so that some of the wines didn't pair well with some of the dishes, he would continue to behave as he does, resulting in bad recommendations. As

¹⁰¹ My definition of guidance is a development of Railton's (2006) with which it is helpful to compare it:

Agent's A conduct C is guided by norm N only if C is a manifestation of A 's disposition to act in a way conducive to compliance with N , such that N plays a regulative role in A 's C -ing, where this involves some disposition on A 's part to notice failures to comply with N , to feel discomfort when this occurs, and to exert effort to establish conformity with N even when the departure from N is unsanctioned and non-consequential. (p.13)

There are three central differences between my definition and Railton's: (i) I understand the characteristic regulative function of guidance as irreducibly normative, whereas Railton seems to suppose that it can be reduced to dispositions that, I take it, are not normative on his view; (ii) relatedly, these dispositions in turn require the agent to have some awareness of her failing or succeeding at the task, whereas my definition requires no such awareness; and (iii) I characterize guidance in terms of an aim rather than a norm.

such, his performance is not guided by the aim of providing good recommendations. By contrast, the recommendations of the waiter who has the relevant knowledge or intuition meet both (i) and (ii). Hence, they are competently performed.

Epistemic performances, we are assuming, aim at truth and the beliefs that they give rise to are warranted just in case they are competently done relative to this aim. Putting it all together, then, the proposal is this:

Warrant Through Competence (WTC): *S* is warranted in forming/maintaining the belief that *p* iff (a) *S*'s belief is formed/maintained through a performance that is reliable at ensuring that *p* is true, and (b) it is so reliable in virtue of being guided by the aim of ensuring that *p* is true.

In the previous section, we saw that classical reliabilism yields false positives. For instance, it incorrectly predicts that in **Brain Lesion**, Braid knows that he has a brain lesion. We can avoid this result by appeal to (WTC). Although Braid's belief meets the reliability condition (a), it doesn't meet the guidance condition (b). Indeed, it falls short of meeting (b) in two respects. First, brain lesions do not have an epistemic aim, neither as such nor as part of some further performance. Second, and a fortiori, a belief formed on the basis of a brain lesion cannot be formed through a performance that is reliable in virtue of having truth as its aim. For the performance as a whole, insofar as there is one, does not even have truth as its aim.

The fact that it avoids such false positives is a central advantage of the proposal. I also suggested that we can test theories of direct warrant in terms of how well they account for perceptual knowledge. Let us now consider how the account fares in this regard.

When we form beliefs on the basis of perception we do so on the basis of an information-gathering and -delivering system with a particular structure. It is helpful to make this structure clearer by distinguishing between three tasks (or sub-performances):

1. *Information-check*: A system ensures that certain information holds true of the world.
2. *Delivery*: The information is delivered in an appropriate form to be used at a next stage of the process.
3. *Dependent Task*: A performance of a determinate sort is taken on the basis of the information thus delivered.

Each of these tasks is a performance in its own right, with its own determinate sub-aim; but together they might constitute a broader performance with an overall aim to which each subtask contributes. In the case of perception, our perceptual systems perform a series of complex tasks to check information concerning our surroundings: for instance, in the case of vision, some movement might direct my attention to some movement in my garden, eventually registering that there is a squirrel running through it.¹⁰² Second, the information is delivered in an appropriate form to be assessable for belief-formation: in the case of perception this can happen through conscious perceptual awareness of the squirrel as it runs.¹⁰³ Third, when this information is thus available, I am in a position to form a perceptual belief: on the basis of the perceptual awareness of the squirrel, I form the belief that a squirrel is running in my garden.

Given that a process has such a tripartite structure, what does it take for it to be competent by the lights of the account of competence I have provided? I suggest that to answer this question we should appeal to the following principle:

Competence by Parts (CP): Given a complex performance p with an overall aim A that is constituted by a series of sub-tasks $\langle t_1, t_2, \dots, t_n \rangle$, p is competently

¹⁰² Of course, this describes only a small part of the process, and at a high level of description. A lower-level description will include the adjustments of the retina in responses to changes in the lighting conditions, and such. I say ‘check’ to emphasize the progressive nature of this activity, with continuous adjustment to achieve the aim.

¹⁰³ But it need not: in cases of blindsight, information is available for doxastic consideration without consciousness.

performed (by the standards of *A*) iff the subtasks in the series, $\langle t_1, t_2, \dots, t_n \rangle$, are competently performed.

I take this to be a fairly intuitive principle. Essentially, it says that the competency of a complex performance supervenes on the competency of the parts.¹⁰⁴ Once again, though, one of the advantages of working from within a virtue-theoretic framework is that we can test these principles in a wider context, considering how they fair when applied to non-epistemic performances.

The case of perception is an intra-personal, epistemic performance; but there are inter-personal, non-epistemic performances with the same tripartite structure. For instance, consider the standard operation in a restaurant of a chef and a waiter. We can distinguish the following three subtasks: (1) the server's taking down the order aiming to take down what the customer wants; (2) the deliverance of the order to the kitchen; and (3) the dependent task, in this case the chef's cooking of the meal on the basis of the order that the server delivers. These three subtasks are all part of a broader task whose aim at having the right meal cooked.¹⁰⁵

Suppose the waiter delivers an order to the kitchen. It reads, 'Pasta Arrabiata'. What conditions must be satisfied for the performance as a whole, culminating in a certain dish, to be competent? Appealing to (CP) I answer that the performance is competent just in case (1), (2), and (3) are themselves competently performed. This seems to yield the right results: if the waiter competently takes the order, competently delivers it to the kitchen,

¹⁰⁴ The parts need not be assumed to be capable of independent specification apart from the complex performance or overall aim.

¹⁰⁵ A fuller description of the overall task would involve the server's bringing the food to the customer, but we can ignore that part of the performance for our purposes.

and the chef competently cooks it on that basis, then it seems as though, if the right meal is cooked as a result, the result will be a competently cooked meal.

Notice a crucial implication of this account: it entails that at no stage in the process need there be any checking or supervision of the waiter's part on behalf of the chef, who can simply *rely* on the waiter's performance. This corresponds to a non-internalist condition for epistemic performances: when we form beliefs on the basis of certain systems, we need not supervise the operation of those systems. We can simply rely on those systems. This is clearly the right verdict in the case of the restaurant. Indeed, in ordinary cases, any further checking would manifest lack of competence on the part of the chef: all things equal, competent cooks should rely on their waiting staff—doing more amounts to inept micromanagement.¹⁰⁶

Returning now to the case of perception, what needs to be shown is that each of the three stages is competently performed (at least normally). That this is so, is clear. First, our perceptual systems are designed to register true information about the world, and they adjust to ensure that only true information is registered. This is done in complex and varied ways, from adjustments to the retina in response to changes in lightning conditions, to changes in what is perceived in response to changing circumstances (e.g. if the squirrel stops moving, it will be so registered). This information is competently made available for doxastic consideration through consciousness, for instance. If the squirrel is moving, I will be perceptually conscious that it is doing so; and I will be conscious of its stopping when

¹⁰⁶ Although I focus on a case where the tasks are performed by distinct persons, this is incidental to the point. The same account would apply to a case where a machine (or any such instrument) takes the order and delivers it. Thanks to Steve Darwall for encouraging me to address this. I should also note that the fact that excessive care manifests incompetence in some cases, is problematic for reductive versions of virtue reliabilism. This is because excessive care does not as such prevent the acquisition of knowledge. Hence, we cannot define epistemic warrant by appeal to any old competence; it has to be competence of the right kind.

it does. Finally, people are normally competent at forming beliefs on the basis of the information delivered by their perceptual system. This includes the skill of discerning in which cases these systems may not be so trustworthy; but quite generally it simply consists on believing what one perceives. In virtue of all this, our perceptual systems reliably yield true beliefs. By the lights of the account presented above, this means that they are warranted. In the right environment, they therefore constitute knowledge.

I would like to note one further advantage of the account of epistemic warrant defended here. Virtue-epistemic accounts of the sort presented here are sometimes accused of conceptual circularity.¹⁰⁷ On these accounts, whether a subject has knowledge depends on whether she forms beliefs competently. However, the objection goes, competence requires knowledge. Hence, we need to appeal to competence to define knowledge, but we also need to appeal to knowledge to define competence. The account I have defended is not subject to this criticism. Since the definition of competence that I have presented does not appeal to knowledge, we can avoid the charge of circularity.

3. CONCLUSION

In this chapter, I have defended a version of virtue reliabilism by showing that it avoids the false positives of classical reliabilism, while still explaining how it is that we acquire knowledge directly on the basis of perception. It should be clear that the kind of explanation that was given of how we acquire perceptual knowledge could also be given for how we acquire knowledge on the basis of memory and inference, for instance. As

¹⁰⁷ See, e.g. Stanley and Williamson (2016) for this charge.

such, it provides a good basis to develop a Reidean epistemology that admits various sources of direct knowledge. I shall thus appeal to this account in the next two chapters, as part of an argument to the effect that the Reidean list should be expanded to include practical sources.

Chapter 5: The Will as Epistemic Faculty

Appealing to the account of direct warrant developed in the previous chapter, this chapter presents my positive account of agential knowledge. This account is designed to meet the four desiderata identified at the end of chapter 3: it aims to show (1) why there is a non-accidental connection between intentional action and agential knowledge, (2) why expressions of intention naturally and correctly tend to take assertoric form, (3) that agential knowledge is direct, and (4) that agential knowledge is spontaneous.

The account consists of two central theses. The first is a thesis in the metaphysics of mind. I argue that the intentions that constitute agential knowledge are beliefs. This can sound much like the traditional cognitivist claim that intentions are beliefs, a claim that I rejected in Chapter 2. However, this claim is traditionally understood as a definitional claim, the claim that what it is to be an intention is to be a belief. I reject this view, since it implies that one cannot have an intention without having a corresponding belief about one's actions. However, there is a generic reading of the claim that I endorse: in general (but not always) intentions are beliefs. According to this view, the token mental states that constitute intentions are, in many cases, beliefs. This is the view I shall be developing in the first part of this chapter (Section 1). As we shall see, however, holding this view seems to lead to epistemological problems: it seems to commit us either to a pragmatist view on which practical reasons are reasons for belief, or to the view that the beliefs that are intentions could not be epistemically warranted. The second thesis aims to avoid this problem. It is the view that intentions that are beliefs are epistemically self-warranted. Qua

intentions they provide epistemic warrant for themselves qua beliefs. I defend this view by appeal to the account of warrant developed in the previous chapter (Section 2).

1. A MODERATE (TOKEN-LEVEL) COGNITIVISM

Recall that traditional cognitivism is often defended by appeal to the fact that intentions are most naturally expressed in the form of flat-out assertions (e.g. ‘I shall take a walk’). Cognitivists often take this to mean that they express beliefs. As noted, however, even if this step of the argument is sound, it falls short of supporting traditional cognitivism. The problem is that there are cases where an agent intends to do something but is not in a position to assert outright that she will do it, occasions where it seems natural to say instead, ‘I am trying to Φ ’ or ‘I intend to Φ ’ but not, ‘I am Φ ing’ or ‘I will Φ ’.¹⁰⁸ These are precisely the cases that call into question principles like (IEB) and (AEC), where the agent intends to do something but believes she won’t do it. The carbon-copier, for instance, seems not to be in a position to assert, ‘I will make ten carbon copies’, precisely because she is unsure whether she will.

What needs to be explained, therefore, is why *normally but not always* intentions are naturally and properly expressed in assertoric form. The cases where they are not so naturally asserted call traditional cognitivism into question. I thus suggest that the argument from assertoric expressions of intention supports a weaker thesis than traditional cognitivism, namely, the view that *some* intentions are beliefs. Such a view allows us to hold that in the cases where intentions are naturally expressed as assertions, those

¹⁰⁸ Building on work by Ferrero (2009), Marušić and Schwenkler (2018) present an alternative explanation of these assertions in terms of partial intentions. However, I take the argument against (AEC) above to rule out even this possibility, and the explanation I shall offer is in any case simpler.

intentions are beliefs; but in the cases where one is at most willing to assert that one intends to do it, one's intentions are not beliefs. I shall refer to this view as 'moderate cognitivism'.

To make sense of the proposed view, consider a case where a speaker performs more than one illocutionary act with a single utterance, as in Anscombe's example near the beginning of *Intention*: "when a doctor says to a patient in the presence of a nurse 'Nurse will take you to the operating theatre', this may function both as an expression of his intention . . . and as an order, as well as being information to the patient" (1958: 3). For simplicity, let us focus on the act insofar as it is both a *command*, a speech-act whose function is roughly to make a subordinate do something, and an *assertion*, a speech-act whose function is roughly to inform the listener of some fact. On the view I am proposing, just as a single utterance can be both a command and an assertion, a single mental state can be both an intention and a belief; 'intention' and 'belief' are thus sortal terms on this view, and the contention is that there are token mental states that fall under both sorts. This contention does not imply that all beliefs are intentions or that all intentions are beliefs, anymore than the contention that some commands are assertions entails that all commands are assertions or that all assertions are commands. As such, it does not have the problematic implications of traditional cognitivism: moderate cognitivism is compatible with the view that you can intend to do something without having any positive cognitive attitude about your intended action.

How one understands the claim that a single mental state is both a belief and an intention will depend on one's favoured theory of mind. I shall concentrate on a kind of dispositional account that I find attractive, according to which a mental state is of a certain type in virtue of having a certain dispositional profile. However, the considerations I appeal

to should make it possible to see how the claim would be understood from the standpoint of various other theories of mind.¹⁰⁹

Here are two dispositions that are plausibly taken as constitutive of the dispositional profile for belief (I use ‘ $s(p)$ ’ to stand for ‘a state s with content p ’):¹¹⁰

(1a) *Evidence-responsiveness*: $S(p)$ is a belief of A ’s only if A is disposed to give up $s(p)$ in response to evidence to the effect that p is false;

(2a) *Input for theoretical reasoning*: $S(p)$ is a belief of A ’s only if A is disposed to use p as a premise in her theoretical reasoning.

Corresponding dispositions yield a parallel dispositional account of intention:

(1b) *Action-governance*: $S(p)$ is an intention of A ’s only if A is disposed to make it the case that p ;

(2b) *Input for practical reasoning*: $S(p)$ is an intention of A ’s only if A is disposed to use p as a premise in her practical reasoning.

¹⁰⁹ For instance, it is easy to see that one could make sense of moderate cognitivism along (a) interpretationalist, (b) inferentialist, or (c) functionalist lines. The reason is that the dispositions I focus on are the sort that (a) are likely to be used as a basis for the relevant interpretation, (b) concern inference-licencing, and (c) are plausibly constitutive of the functional profile that characterizes belief and intention. It should also be evident, given the kinds of dispositions I go on to focus on, that the account is not intended to be reductionist, neither of the specific mental states in question (belief and intention), nor of mental states in general, nor of the notion of content.

¹¹⁰ For simplicity, I shall take the content of intentions to be propositions. This view has been contested on the basis of (i) the de se nature of this content (see e.g. Thompson (2008: 120-27)) and (ii) the infinitival structure of its expression (see e.g. Hornsby (2016) who extends the objection to intellectualist accounts of know-how). The first charge seems to me to depend on an overly restrictive account of propositions that does not allow for de se propositions (as found e.g. in Lewis (1979)). Such a restrictive account has the implausible implication that the contents of the belief that I was at the library yesterday and my belief that someone was at some library on Tuesday are of a fundamentally different type (or else that the latter belief is also non-propositional, if we follow Lewis). On the second point, the extended discussion in Stanley (2011, ch.3-5) leads me to side with the propositionalist on this issue.

An alternative route would be to identify a certain class of beliefs with the alleged non-propositional type of content that intentions have and then identify intentions with such beliefs. Perhaps beliefs corresponding to the speech act of claiming-to, as opposed to asserting-that (e.g. I claim to be a good cook/be at home/be running) would work. If the anti-propositionalist can make sense of such beliefs, her view is compatible with my proposal. For the record, though, I am inclined to think that claiming-to ascriptions are also (de se) propositional. Ultimately, my central commitment is to the view that intentions are about “what happens” in Anscombe’s (1958) sense (p.52): the kind of thing over which there might be “knowledge, or opinion based on observation, inference, hearsay, superstition or anything that knowledge or opinion ever are based on” (p.50). See McDowell (2010) for a convincing defense of this view. Although his argument focuses on knowledge of occurrent action, the considerations he appeals to seem to generalize for knowledge of future action (absent, of course, a general skepticism about knowledge of the future, a view I find independently implausible).

The dispositions outlined so far are descriptive in nature but, in my view, a dispositional account will yield an impoverished view of belief and intention unless it recognizes the normative dimension of these mental states. For I take it as a central point of convergence of much recent work in philosophy of mind that part of what it is to have these mental states is to be subject to certain normative constraints.¹¹¹ Beliefs, for instance, are mental states that are normatively (and not just dispositionally) regulated by truth, as captured by the following norm:¹¹²

Truth Norm for Belief (TNB): Believe p only if p is true.

In turn, this gives rise to *rational* dispositions, dispositions to regulate one's beliefs according to this norm, yielding:

(3a) $S(p)$ is a belief of A 's only if A is disposed to regulate $s(p)$ according to whether it conforms to TNB.

Similarly, on an influential view (Bratman 1999 [1987]), to have an intention is to have a mental state that is subject to the following norms:¹¹³

Consistency Norm for Intention (CNI): Intend to Φ only if Φ ing does not conflict with your other intentions.

¹¹¹ See, among many others, e.g. Bratman (1999 [1987]); Hieronymi (2006); Shah (2003, 2008); Shah and Velleman (2005); Wedgwood (2013a, 2013c, 2013b, 2002); Williamson (2000).

¹¹² In fact, I am inclined to follow the suggestion in Williamson (2000) that the norm of belief should be a knowledge- rather than a truth-norm. For simplicity, though, I focus on the weaker truth norm.

¹¹³ There is presumably a corresponding coherence norm for beliefs, but I would regard it as derivable from (TNB) or a knowledge norm. In fact, I am inclined to think that parallel to a knowledge-norm for belief (as per the above fn.) we should recognize an intentional action norm for intention (intend to Φ only if you will have Φ 'd intentionally), since I think parallel arguments support both types of norms, and that we can subsume both of Bratman's norms under it. I hope to be able to defend these claims in future work. For simplicity, though, I focus on the better-known norms from Bratman. In light of the considerations in Yaffe (2010), I take the requirements arising from these norms to be merely pro tanto. I also assume that the dispositions that they ground (both in the case of belief and intention) are general and fallible: x may be disposed to Φ in circumstances of the C-type, even if there are circumstances of the C-type where x does not Φ (e.g. because the disposition is blocked or masked (Johnston 1992)). Thus, it is sufficient to count as believing or intending something if there is a *general tendency* to conform to these norms. This is required in any case to account for irrational beliefs.

Inertial Norm for Intention (INI): If you intend to Φ , do not easily reconsider (in deliberation) whether to Φ .

In turn, this gives rise to the following rational disposition:

(3b) $S(p)$ is an intention of A 's only if A is disposed to regulate $s(p)$ according to whether it conforms to CNI and INI.

Finally, it is worth noting that there might well be important relations between these various dispositions. For instance, it is attractive to hold that part of the reason why it is rational to be disposed to use beliefs and intentions as premises in reasoning (dispositions (2a) and (2b)) is precisely that they conform to their respective rational norms.¹¹⁴

Imagine now a simple dispositional view on which the a- and b-conditions above are not only necessary but jointly sufficient for a mental state to constitute, correspondingly, a belief and an intention. From the standpoint of such a theory, then, what it would take for there to be a state that is both an intention and a belief would be for there to be a representational mental state that displayed dispositions (1a)-(3a) *and* (1b)-(3b) relative to a single content. Take, then, an ordinary intention, such as the intention to run at the gym next morning, and let's assume that it constitutes an intention precisely in virtue of displaying the b-dispositions. Suppose that, after forming this intention, you receive a notification by phone that the gym will be closed for maintenance the next day. You would then have evidence suggesting that what you intend to bring about won't happen, and it is natural under such circumstances to give up one's intention. Indeed, this example illustrates a general dispositional feature of this and many intentions: when agents face overwhelming evidence to the effect that they can't accomplish what they intend, they

¹¹⁴ Inferentialists, of course, might reject this view which gives priority to norms that appeal to alethic notions over the inferences they license.

tend to either revise or give up their intention. This suggests that at least some intentions display (1a) and (3a). Now suppose that you get no such notification, but that your landlord asks you if you will be home the next morning. In considering this question it would be natural for you take for granted the content of your intention (to wit, that you will be at the gym the next day), and therefore to tell the landlord that you won't be at home all morning. Once more, this illustrates a general feature of this and many intentions. Hence, at least some intentions meet condition (2a).

On the other hand, not all intentions display the dispositions I have identified as constituents of beliefs. First, one might be so committed to a certain course of action that one would not give up one's plan regardless of how impossible it may seem. An intention that displayed such a commitment might not meet (1a) or (3a). Second, if one is unsure about whether one will accomplish what one intends (e.g. one knows that one is weak-willed when it comes to going to the gym), one should not rely on its content in theoretical reasoning.¹¹⁵ Since we are taking (2a) as necessary for a mental state to constitute a belief, this means that such intentions do not constitute beliefs. Given a dispositional account along the lines suggested, therefore, we can easily make sense of how there could be some intentions that are beliefs, without holding that all intentions are beliefs.

There are, however, two important objections to consider. First, one might suggest that once a dispositional theory is fully spelled out, the dispositions associated with an intention will be strong enough to ensure that a mental state that displays all of them

¹¹⁵ Thanks to David Charles for pointing the relevance of weakness of will for this discussion. I regard it as a central advantage of the proposal that it can account for such cases naturally. I hope to explore this upshot in future work.

constitutes a belief.¹¹⁶ Yet, I believe that the objections against traditional cognitivism in Chapter 2 rule out this possibility. These objections suggest that one can have an intention without having *any* theoretical commitment towards what one intends. Hence, they suggest that one can have an intention without having a belief of the relevant kind, and a dispositional theory that cannot account for this should on this basis be rejected. Any dispositional account of these mental states must, like the theory briefly outlined above, ensure that some intentions are not beliefs.

The second strategy is to suggest that beliefs and intentions have incompatible constitutive dispositions, so that no mental state could be of both types.¹¹⁷ For instance, one might argue that, despite what I have said, intentions are strongly *irresponsive* to evidence, so that they fail to display (1a). Someone may intend to do something despite strong evidence to suggest that they won't accomplish what they intend: for instance, someone can intend to quit smoking despite strong evidence to suggest that she won't quit.

I have two replies. First, we should not make (1a) overly strong. After all, it is clear that many beliefs are strongly *irresponsive* to the evidence; these beliefs may be irrational, but they are beliefs. Second, and most importantly, whether a consideration counts as evidence for a person depends on the epistemic position of that person; and this position is altered in particular by how she antecedently relates to the matter that the purported evidence bears on. To see this, consider the following example from Rowe (1979: 340). A ship has suffered a terrible accident of a sort that makes it overwhelmingly likely that no

¹¹⁶ I thank Beri Marušić for pressing this objection on me. Note that even though I keep focusing on dispositional accounts, it is easy to present the objections in other frameworks (see fn.33).

¹¹⁷ See Smith (1994: 118) for an influential line of objection along these lines, and Setiya (2007) for what I take as a convincing reply to the specific version of the problem as presented by Smith. I thank David Charles and Gideon Yaffe for pushing an objection of the sort considered below.

one in the ship survived. On the basis of this evidence, it would be wrong for most people to believe that anyone survived. However, suppose you were on the ship, and you miraculously survived. You would then be aware of your survival, in which case you should not give weight to considerations that suggest that no one survived: your awareness of surviving should bracket other considerations that would tell against your surviving. I suggest that the same holds for the beliefs that an agent normally forms about her actions: because these beliefs concern what is in her power to bring about, certain forms of evidence, such as statistical evidence, do not weigh against them.¹¹⁸

On the basis of the considerations above, I conclude that moderate cognitivism seems to be a coherent position. Of course, a dispositional account along the lines presented above would need to be refined and defended if it is to be used to *support* moderate cognitivism; but my purpose in bringing it up here is not this, but instead the more modest one of providing a way of making sense of the view. Granting that the view is coherent, we have already on the table what I take to be powerful arguments in its favour. For appealing to this view enables us to meet two of the desiderata for a theory of agential knowledge.

First, the view seems to best account for the fact that intentions are most naturally expressed as outright assertions, but also for why they are not always naturally asserted thus. When an intention is a belief, it is natural and appropriate (if it constitutes knowledge) to express it as a flat-out assertion. However, when the intention is not a belief, this is

¹¹⁸ See Marušić (2012) for an extended treatment of this topic and how something's being up to us can alter our epistemic relation to some matter. His view is that beliefs about our actions are subject to different, non-evidential norms. This would be an alternative way to defend moderate cognitivism against the present objection.

neither natural nor appropriate. Yet, the person might know that they have an intention to do something. In that case it might be correct for them to assert, 'I intend to do such and such' even if they are not in a position to say, 'I shall do such and such'.

Second, the view has the resources to explain the spontaneity characteristic of agential knowledge. Because intentions in general are states that spontaneously settle how the world is (or will be), beliefs that are intentions are spontaneous in the required way. Agential knowledge is constituted by beliefs that are intentions, and this explains why it is spontaneous. On the other hand, since moderate cognitivism does not require us to posit implausible necessary connections between cognitive and practical entities, the view is capable of explaining this feature of agential knowledge without appeal to implausible metaphysical connections, as required by traditional cognitivism.

On the other hand, it might appear as though cognitivist views, including moderate cognitivism, are particularly *unsuited* to explain how it is that the relation of an agent to her actions is one of knowledge. To see this, consider the case where I have formed an intention to go running in the afternoon because it is healthy and enjoyable. That running is healthy and enjoyable are good reasons to form (and maintain) this intention. However, they do not appear to be good epistemic reasons, reasons to hold the belief: these reasons speak for the worthiness of the intended action, but holding a belief just on the basis of such reasons would seem to amount to a problematic form of wishful thinking. However, such wishful thinking cannot amount to knowledge. Therefore, even if we could make

sense of an intention that is also a belief along the lines suggested, such an account could not explain how we ever come to have practical knowledge.¹¹⁹

The problem can be presented in terms of a more general argument as follows:

1. Intentions are held on the basis of practical reasons.
2. Beliefs held on the basis of practical reasons are not in good epistemic standing. [Assumed]
3. If a belief constitutes knowledge, then it is in good epistemic standing. [Assumed]
4. Take some arbitrary mental state $S(p)$ that is both an intention and a belief. [Assumed for Universal Generalization]
5. Therefore, $S(p)$ is held on the basis of practical reasons. [1 and 4]
6. Therefore, $S(p)$ is not in good epistemic standing. [2 and 5]
7. Therefore, $S(p)$ does not constitute knowledge. [3 and 6]
8. Therefore, mental states that are both intentions and beliefs cannot constitute knowledge. [4 and 7]

I take 1 and 3 to be uncontroversial. Many cognitivists, however, dispute premise 2. They argue that the beliefs that constitute practical knowledge are special precisely because they are justified by practical reasons, a case of “licensed wishful thinking” as Grice (1971) puts it. In this vein, Marušić and Schwenkler (2018) argue as follows:¹²⁰

One may also hold a practical belief, therein accepting something so as to create truth, not in light of evidence that this is going to happen, but on the basis of considerations that show this truth to be worth creating. Our position is that intentions are beliefs of this second sort: An intention is a commitment to make something true, in light of the fact that this truth appears worth creating. Thus we hold that practical reasoning does support the conclusion that something is (or will be) true. (p.315).

Although I sympathize with the claim that practical beliefs (those that are constitutive of agential knowledge) aim at truth insofar as they aim to bring it about, I find it implausible

¹¹⁹ See Langton (2004) for an extended objection to cognitivism along these lines. In her (2009), however, Langton argues that a broadly externalist outlook can be used to explain how ‘practical beliefs’ can be knowledge. The argument in the next section is congenial to this claim.

¹²⁰ See also Velleman (1989, p.129 et passim) for similar views. However, I take Velleman to hold ultimately that practical knowledge is warranted by the evidence that the agent has that she will act based on what she desires most (which generally will go with what she thinks she has reason to desire). Thus stated, though, the view is subject to a version of the argument against inferentialism in Section 1.

to hold that practical reasons provide epistemic warrant. Let us grant that there are cases where practical reasons speak in favour of forming some belief or other. Even if such reasons make forming a certain belief all things considered correct, it is hard to see how such a belief could be made *epistemically* correct by those reasons. Thus, imagine a variant of William Tell's story, where an *unskilled* archer is forced to shoot an arrow towards his son, who has an apple on his head:¹²¹ if he hits the apple, he and his son are free to go; but if he does not shoot, he will be killed along with the rest of his family. The archer thus has excellent practical reasons to intend to hit the apple. He might also believe (correctly, let's assume) that if he believes that he will hit the apple, his chances of hitting it will improve. Thus, he has excellent practical reasons to believe that he will hit the apple and to make this true. Yet, if he so believed, the belief would clearly be badly formed by epistemic standards. Thus, it seems that even excellent practical reasons do not provide epistemic support for beliefs.¹²²

In my view, then, we should endorse premises 1-3. However, I reject the conclusion, because I also hold that the argument is invalid. It depends on the false assumption that reasons-attributions are extensional. But they are not: they are description-dependent. This means that the transition from 5 to 6 is illicit. Let me explain.

The transition from 5 to 6 is an instance of a more general inference pattern with the following form:

¹²¹ We may assume that though unskilled, the archer knows how to shoot an arrow towards a target.

¹²² I have set up the case so that the agent has both strong reasons to act (and, so, to intend) as well as reasons to form the intention, if we distinguish the two, as I think we should to account for Kavka's paradox. What the case shows is that the strength of reasons of either kind does not as such contribute to improving a subject's epistemic standing, and cannot therefore be the proper basis for belief.

9. x is formed for reasons R_1, R_2, \dots, R_n .

10. $x=y$.

11. Therefore, y is formed for reasons R_1, R_2, \dots, R_n . (9 and 10)

If an inference of this form is invalid, we cannot move from the claim that the intention is formed for certain reasons to the claim that the belief with which it is identical is formed for those reasons. On the other hand, the inference seems to depend on a harmless substitution of co-referring terms. Of course, whether the substitution is harmless depends on whether 9 is extensional in the x -position (henceforth the ‘supported position’). This is what I wish to dispute.

Consider the following case:

A Quick and Loud Lecture: Mercedes has accepted an invitation to present a public lecture in a park. The audience is quite large, so to allow her voice to reach as many people as possible, Mercedes speaks loudly throughout the whole presentation. Right before beginning the lecture, however, Mercedes also realizes that she must finish the lecture in 2 hours to be able to pick up her child from daycare. Thus, Mercedes speaks quickly throughout the whole presentation.

In this case, it seems true to say that:¹²³

- 12. The reason Mercedes is speaking loudly is to be heard by a large audience;
and
- 13. The reason Mercedes is speaking quickly is to be able to pick up her child from daycare.

Yet, we may assume that there is a single speaking event (or process) going on, a speaking that is both loud and quick. In other words:

14. Mercedes’s speaking loudly=Mercedes’s speaking quickly.

Suppose, then, that substitution of co-referring terms was permissible in the supported position. We could then infer that:

¹²³ Although I focus on sentences of the form ‘The reason that S Φ s is x ’ note that the point holds for reasons-attributing sentences generally (e.g. ‘ S Φ d because x ’, ‘ S Φ s in order to x ’, etc.)

- 15. The reason Mercedes is speaking quickly is so as to be heard by a large audience; and
- 16. The reason Mercedes is speaking loudly is so as to be able to pick up her child from daycare.

However, (4) and (5) are evidently false. Hence, we should reject the assumption that reasons attributions are extensional in the supported position: *R* is a reason for *x* only under particular descriptions of *x*.¹²⁴

The argument could be resisted by holding a so-called ‘fine-grained’ account of action individuation. Take, for instance, Alvin Goldman’s (1970) view that action-tokens are individuated in terms of action-types or “act-properties” (p.10ff.).¹²⁵ Since the act-properties of speaking loudly and of speaking quickly are different, it follows from Goldman’s account that Mercedes’s loud speaking and Mercedes’s quick speaking are different actions. This means that 14 is false, and that is why we obtain falsehoods 15 and 16, rather than, as I have suggested, because of the intensionality of reasons-attributions.

However, a fine-grained account of this sort has unacceptable implications. Not only does it entail that there were several more actions where there is intuitively only one (e.g. a loud speaking, a speedy speaking, a clever speaking, etc., where intuitively there is only one act of speaking with all these properties), but it also conflicts with seemingly innocuous inference patterns.¹²⁶ To see this, consider the following inference pattern:

- 17.
 - a. $G(\iota x)(Fx)$
 -
 - b. $(\iota x)(Fx) = (\iota x)(Fx \ \& \ Gx)$

¹²⁴ The converse, that *R* is a reason only under certain descriptions of it is also true, and widely endorsed.

¹²⁵ A similar view is found in Rescher (1969) and receives sustained attention as a general account of event-individuation in the writings of Jaegwon Kim (1973, 1976, 1966).

¹²⁶ This is a point made by Katz (1978) against Kim’s account, but it generalizes to all fine-grained accounts.

The inference pattern seems valid, as suggested by the following instantiation of it:

18.
 - a. The table in the room is brown.

 - b. The table in the room=The brown table in the room.

But now consider a straightforward application of 17 to actions:

19.
 - a. Mercedes' loud talk was quick.

 - b. Mercedes' loud talk = Mercedes' quick talk.

Since the fine-grained theorist denies 19b, she must either (i) reject the validity of the inference pattern in 17, or (ii) reject premise 19a. In the absence of further argument, (i) is ad hoc, while (ii) just seems outrageous: if Mercedes's loud talk was not quick, then at what speed was it delivered? (After all, it must have been given at some speed!)

Although a full investigation of the topic is not possible, I take such arguments to show that there is strong pressure to accept a coarse-grained account of action individuation.¹²⁷ As such, I accept 14, and conclude on the basis of the foregoing argument that reasons attributions are intensional.

Thus, we can hold that the beliefs that constitute agential knowledge are intentions, while denying that the practical reasons that warrant holding such mental states qua intentions also warrant holding such mental states qua beliefs. On the other hand, we seem to be out of the frying pan and into the fire: if the beliefs that constitute agential knowledge

¹²⁷ I discuss fine-grained views at greater length in Piñeros Glasscock (ms), where I argue that one of the central purported advantages of this view—the ability to render causal sentences extensional—is illusory: the view cannot account for the intensionality of many causal sentences, and, in any case, the intensionality is harmless and compatible with the objectivity of causation. At the time of writing, a draft of this paper is available on my website.

are held (as such) neither on the basis of practical reasons nor inferential reasons (as suggested in Chapter 3), then, on what basis are they held? We seem to be out of reasons to appeal to; but saying that these beliefs are held on the basis of no reasons seems hardly better than saying that they are held on the basis of practical reasons if our goal is to show that they are in good epistemic standing.

My broad response to this problem can be anticipated by my response to Agrippa's trilemma. In my view, what we should hold is that the beliefs that constitute agential knowledge are foundational beliefs. I shall argue for this conclusion partly on the basis of the account of warrant developed in Chapter 4 for the case of perceptual knowledge. Yet, there is a crucial difference between the warrant of agential and perceptual beliefs: the former, I shall argue are self-warranting beliefs. I defend this idea in the next section.

2. SELF-WARRANT THROUGH COMPETENCE

In the previous chapter, I took perceptual beliefs as paradigm instances of foundational beliefs. The other type of beliefs that have traditionally been taken to have the right features to be foundational are self-warranted beliefs. Thus, in his *Meditations* (2013 [1641]), Descartes argues that a belief with the following content can put a stop to the skeptical doubts he has raised (p.17):

20. I am thinking.

Now, there are familiar objections to Descartes's contention that this belief is beyond doubt and to the project that he undertakes on the basis of this assumption, objections that I don't intend to discuss. My goal in bringing up the cogito is more modest. For, even if it cannot play the anti-skeptical role that Descartes envisions, he is clearly right to think that each of

us can know that we are thinking, and that we know this in a special way, not available to others.

Descartes's explanation of why this knowledge is available in a distinctive way only to the subject relies on the fact that 20 meets the following condition:

21. **Self-Verification:** Necessarily, if I believe that I am thinking, then I am thinking.

For believing 20 is a form of thinking, so by believing 20 one ensures that the very belief is true: one cannot believe 20 falsely. By contrast, believing that someone else is thinking does not ensure that she is thinking. Thus, the very act of believing that I am thinking provides warrant for the truth of itself.

For Descartes it is crucial that the epistemic warrant conferred to the thinker by holding the belief in 20 is non-inferential. In particular, the thinker need not believe 20 on the basis of his belief in 21; rather, 21 simply explains why a belief in 20 would be warranted if held. It is hard to see, at any rate, how a thinker could believe the antecedent of 21 (that she believes she is thinking) without already believing that she is thinking. And belief in the antecedent would of course be a precondition for inferring 20 on the basis of 21. Thus, if we made belief in 21 a precondition to being warranted in believing 20, we would end in a vicious circle. Moreover, from a historical perspective, it is hard to see how the belief in 20 could play the anti-skeptical role Descartes wants for it if its warrant was inferential. For it would be subject to his worries about mathematical knowledge to the effect that it is always possible to draw the wrong conclusion on the basis of mathematical reasoning (p.14). On the other hand, if a thinker is entitled to believe 20 without basing that belief on 21, these worries about mathematical inference do not call it into doubt. Thus,

I take Descartes's contention to be that simply believing that one is thinking makes it so that one is (epistemically) warranted in holding this very belief. This is what I mean by saying that this belief is self-warranted, and, in light of the aforementioned circularity problem, I shall take it that Descartes is right that the belief that I am thinking is self-warranted in this way.

What I now want to argue is that the beliefs constitutive of agential knowledge have an analogous (though not identical) self-warranting structure: the nature of these mental states qua intentions make it so that a subject who has formed them is in certain cases directly warranted in holding them qua beliefs.¹²⁸ The argument will be in two steps. First, I shall present a general account of epistemic warrant to explain how a belief could be self-warranted in this way. As we shall see, an account that appeals to self-verification alone is deficient in important respects, and we shall need to appeal to the account from the previous chapter to account for self warrant. Second, appealing to this account, I shall argue that (some) intentions provide direct warrant, so that an agent can be warranted in forming a belief about her actions just in virtue of the fact that this belief also constitutes an intention.

2.1 Self Warrant Through Competence

Granting that a thinker is warranted in believing 20 just on the basis of holding that very belief, let us ask: in virtue of what is she so warranted? Descartes's appeal to 21 suggest that he holds a certain form of reliabilism, the view that a belief is warranted just in case it

¹²⁸ Ultimately, I take this to be an Anscombean thesis. For I am in agreement with Thompson (2011) that, "The overarching thesis of *Intention* was that *self-knowledge in this familiar sense* [of oneself qua oneself rather than other] *extends beyond the inner recesses of the mind, beyond the narrowly psychical, and into the things that I am doing*" (p.200, his italics). This is also a central subject of O'Brien's fascinating book on the topic (2007).

is formed through a process that reliably yields true beliefs. Since the process by which Descartes forms the belief in 20 is self-verifying, it is maximally reliable: it produces only true beliefs.

As we saw in the last chapter, even maximal reliability is not sufficient for warrant, and even an infallibly formed belief can fail to constitute knowledge. Consider a subject who thinks the following thought:¹²⁹

22. I am thinking a thought in perfect English.

Like 20, this thought is self-verifying: so long as a thinker believes 22, the belief is true. Yet, not every belief with this content constitutes knowledge. Consider, for instance, someone in the process of learning English, who often makes mistakes in her diction in this language, even while thinking or expressing propositions as simple as 22. Even though such a person would be guaranteed to form true beliefs when she managed to believe 22, this belief is too accidentally true to constitute knowledge. We can reach a similar verdict about 20 by imagining a child in the process of learning concepts like THINKING. Such a child might manage to think the thought in 20, but if she is still making common mistakes in the application of the concepts involved therein, the belief won't constitute knowledge.

¹²⁹ For the argument to follow, we need to assume a super-fine-grained account of the individuation of content that is language-dependent. Thus, I shall take it that 22 has a different content than (22a):

22a. Estoy pensando un pensamiento en ingles perfecto. [A Spanish translation of 22]

Such a fine-grained account of content is not unmotivated. After all, 22 and 22a have different truth-conditions.

We thus reach the same conclusion that we arrived at in the last chapter, that we need a relation stronger than reliability to characterize epistemic warrant. My suggestion was this:

Warrant Through Competence (WTC): *S* is warranted in forming/maintaining the belief that *p* iff (a) *S*'s belief is formed/maintained through a performance that is reliable at ensuring that *p* is true, and (b) it is so reliable in virtue of being guided by the aim of ensuring that *p* is true.

Applying this definition to the cases above delivers the right results. Thus, consider the difference between the thoughts of a novice and a fluent English speaker, when their thoughts have 20 as their content. In both cases, their performances have as part of their aim ensuring that the beliefs they form are true, and this aim determines whether these performances are correct. Moreover, the way in which the truth of these beliefs is ensured is the same in both cases, namely, by the activation of the thinkers' linguistic abilities. However, only the performance of the fluent English speaker is competent. The reliability of her performance is explained by the way in which her thoughts are structured so as to ensure that the intended content is expressed in her thoughts. By contrast, the reliability of the performance of the novice English speaker is explained by the particular content of the thought she managed to express. Hence only the belief of the fluent English speaker is warranted, and constitutes knowledge. For similar reasons, the Cartesian, 'I am thinking' constitutes knowledge only in the mind of someone competent with the concept of THINKING. Yet, competence with this concept is enough to ensure that the belief is warranted, since a subject exercising this competence would competently ensure that her belief is true.

2.2. Practical Warrant

To know what she is thinking, a subject relies on her conceptual competence. To know what she is saying, a speaker relies on her linguistic competence. I shall now suggest that to know what they are doing or what they will do, agents rely on their practical competence.

Let's start with some terminology. I understand the will to be that faculty in virtue of which we intend and act intentionally. An exercise of the will is a performance that, in the successful case, begins with the agent's intending to Φ and culminates in her having Φ 'd intentionally (in the successful case).¹³⁰ As noted in the previous section, I follow Bratman (1999 [1987]) in taking intentions as distinctive types of mental states subject to certain rational norms, including an inertial requirement to the effect that intentions should be resistant to deliberative reconsideration. However, I take this inertial requirement to require a lot more, for reasons internal to Bratman's framework. On this framework, rational requirements for intentions stem from the role they play in the life of planning creatures, in particular, serving as suitable inputs for further practical deliberation. To serve this role, as Bratman argues, intentions should be resistant to renewed deliberation; but that is clearly insufficient: intentions should also persist in the face of relevant obstacles in bringing about the intended action. It is rational to rely on my intention only if I can expect to adjust my actions so that I actually Φ , even in the face of relevant obstacles: if I'm on the way to the store and the street I was planning to take is closed, I should find a different way rather than give up my intention. My intention would not be a suitable input for further deliberation if I were instead disposed to give it up so easily. All things equal, intentions

¹³⁰ I am here influenced by recent work on intentional action that emphasizes their nature as processes that unfold over time (see e.g. Charles (2018), Ferrero (2017), Thompson (2008), Steward (2012), Stout (1997)). A central epistemological advantage of this view is that it allows us to say that the action is already present (in its initial stage) as soon as the agent intends/believes it to obtain. This enables us to sidestep problems pertaining to knowledge of future contingents.

should be implemented, and failure to implement one's intentions without good reason is a failure of (practical) rationality.¹³¹

This means that an exercise of the will goes wrong when it doesn't culminate in successful action. I take this to be the natural view: if you intend to cook risotto but burn the food, your performance has gone wrong (by the lights of your intention);¹³² if you intend to press 'Play' and press 'Stop' instead, the performance is wrong (cf. Anscombe (1958: 57)). These standards hold not just for what the agent does but also for how she does it. Thus, suppose you are playing pool and intend to sink the 4-ball by hitting the 3-ball. You line up your shot but for lack of chalk it goes in a completely different direction, hitting the 5-ball first, but, low and behold, you hit the 4-ball and sink it in the intended pocket. Here, you achieve the intended outcome, namely, sinking the 4-ball; but the shot was incorrect, precisely insofar as it didn't conform to the way in which you intended the performance to go. Like promises, then, intentions set their own standards for correctness: the performance that begins with an intention to Φ is correct only if the agent goes on to successfully Φ as she intends.

Exercises of the will not only set their own standards, but those standards guide their development which are structured to adjust so as to achieve the aim that the agent sets

¹³¹ Bratman recognizes this feature of intentions, but he characterizes it in terms of a descriptive principle of control, to the effect that a present intention to Φ "will normally lead me to at least try to Φ " (p.108). That this is too weak (both descriptively and normatively) can be seen from the fact that in the example under consideration, I would meet this requirement by my stepping out of the house, even if I gave up as soon as I saw that the street was closed. On failure of implementation as a rational failing, see Holton (1999).

¹³² The qualification in brackets is needed because I am inclined to think that this requirement takes wide scope, so that you could act correctly, in accordance with the requirement, by giving up the intention. Note, however, that unless there are good reasons for giving up the intention, this also falls afoul of the inertial requirement. It should also be clear that the notion of correctness at play is neither moral nor all things considered; it can sometimes be correct all things considered to err in these ways.

for herself through an intention. This can happen in a variety of ways, as illustrated by the following case:

Wedding: In January, I receive an invitation to my friend's wedding on August 19th, which leads me to form the intention to go. I start planning accordingly, booking the hotel, and ensuring that a friend can drive me, etc. I realize that to be able to go I will have to save up and decline otherwise tempting invitations.¹³³ Thus, when I am invited to a costly concert by one of my favourite bands the week before the wedding, I refrain from considering this as an option, and decline to go. It is the 17th and my friend who was supposed to drive me to the wedding calls from the hospital to tell me that he has just been diagnosed with appendicitis, and cannot drive me as a result. I decide to rent a car, look up the instructions, and head out the next day. Unfortunately, through unexpected flooding, the route I picked has been closed. I use Google Maps to find a new route, but I miss one of the exits, so I have to stop again to find my way. Eventually I arrive to the wedding.

The example illustrates Bratman's requirement that intentions should not be too easily reconsidered when I resist the inclination to go to the concert. But it also illustrates the way in which, as they unfold, exercises of will adjust in myriad ways to ensure that one's aim is achieved. After I form the intention, I start taking steps towards carrying out my intended action (booking my hotel, asking my friend to drive me, etc.). When I find out that my friend can't drive me, I rent a car to drive myself; and when I find out that the route I intended to take is closed, I find an alternative one. Finally, I also correct for mistakes I make: when I miss the exit, I find a way to get back on track. Moreover, the case illustrates features characteristic of action execution. The disposition to endeavour in these ways to achieve our intended aim in the face of difficulties and temptations is part of what is distinctive of intentional action as such. Exercises of the will, therefore, are functionally structured so as to ensure that one acts as one intends.

¹³³ See Holton (2009) on the relation between intention and temptation.

Near the end of *Intention*, Anscombe writes: “Surprising as it may seem, the failure to execute intentions is necessarily the rare exception”, but she qualifies the claim a few lines down: “What is necessarily the rare exception is for a man’s performance in its more immediate descriptions not to be what he supposes” (p.87). She qualifies the claim because, as she notes, we often fail to fulfil our more remote ends: I intended to finish a paper last month (alas! I got busy and I’m still working on it); I intended to pick up the book (alas! I forgot to pass by the library); I intended to make 10 carbon copies (alas! I didn’t press strongly enough). Moreover, an agent will reliably achieve her goals only if she is competent at achieving them. Notice that a qualification like Anscombe’s is also needed for perceptual knowledge: although the world is by and large such as we perceive it to be, there are cases where our perceptual systems misfire, for instance, when the lighting conditions are bad, when an object is too far away, or when there is a decoy. Still, there is a large set of cases where our perceptions reliably report facts about the world, and a large set of action-descriptions for which an intention to do them generally culminates in one’s doing them. These descriptions include one’s bodily movements, but also more complex actions like *making a coffee*, *cleaning the floor*, or *going to a wedding*; and the list expands as we gain practical competence. Moreover, it is clear that the reason we reliably succeed in carrying out those actions is precisely that exercises of the will adjust in the way just described so as to bring them about.

To sum up, there is a core set of action descriptions for which it is true that, if one sets about doing a certain action (by forming an intention to so do it), one will reliably do so. And the fact that exercises of will are thus reliable is explained by the fact that they adjust so as to ensure that their contents are true, the aim by which we assess them for

correctness. Exercises of will, therefore, (at least sometimes) competently ensure that their contents are true. Appealing to the account of epistemic warrant developed in the previous section, we reach the result that when we form a true belief as part of performance that is an exercise of will, the belief will be competently formed. Hence it will be warranted.¹³⁴

Admittedly, the way in which exercises of will ensure that their contents are true is different from the way that exercises of our perceptual systems ensure this: when we exercise our perceptual systems, these adjust so as to match the world, but when we exercise our wills the world adjusts to match our intentions. This is one of the features that makes it so that the knowledge at play is *practical*, in Anscombe's sense. Whereas in the case of perceptual knowledge a person relies on her perceptual faculties to adjust to the world, in the case of agential knowledge the person relies on herself as agent to adjust the world as she intends it to be. The fact that the account of warrant developed in the previous chapter provides a unifying explanation of both theoretical and practical knowledge is, in my view, one of its central strengths.

We are finally in a position to return to the two epistemic desiderata for an account of agential knowledge: explaining how the beliefs that constitute practical knowledge are beliefs in good epistemic standing, and explaining the sense in which agential knowledge is direct. I have rejected two ways of meeting the first requirement. First, in Chapter 3 I rejected inferentialism because this view was unable to meet the spontaneity requirement. Second, although advocating the view that the beliefs constitutive of practical knowledge

¹³⁴ Notice that this self-warrant is available only to the agent, the one who forms the intention and is capable of executing it. The present account thus explains another central feature of agential knowledge, namely the fact that it is available only to the agent.

are intentions, I rejected the view that practical reasons could serve as the basis for epistemic warrant.

In this section I defended a third alternative: like the belief that ‘I am thinking’, the beliefs about our actions constitutive of agential knowledge provide warrant for themselves. Suppose I believe that I am running, or that I will go for a run in the afternoon. If they are the sorts of beliefs that are constitutive of agential knowledge, these beliefs will be intentions: an intention in action in the first case, and an intention for the future in the second case. Hence, these beliefs will be constitutive of acts of will through which an agent can competently ensure that what is believed is or will be the case. As such, these beliefs are formed as part of a competent exercise that ensures that they are true. By the theory of warrant offered in this section, these beliefs provide warrant for themselves. Therefore, they are in good epistemic standing, and constitute knowledge. And since their warrant is not derived from other beliefs, the knowledge that we acquire from them is direct rather than inferential.

3. CONCLUSION

A successful account of agential knowledge should meet the following four desiderata: (1) It should explain why we characteristically express intentions through assertions; (2) it should explain the characteristic spontaneity of agential knowledge; (3) it should explain the characteristic directness of agential knowledge; and (4) it should explain how the beliefs that constitute it are in good epistemic standing.

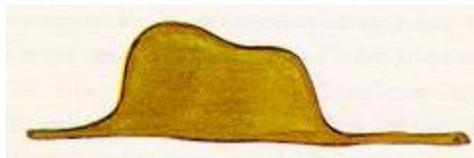
I have defended an account that meets all four desiderata. (1) The intentions that are capable of constituting agential knowledge are also beliefs; such intentions are properly

expressed in assertoric form (when they constitute agential knowledge). (2) Agential knowledge is spontaneous because the beliefs that constitute it are intentions, paradigmatically spontaneous states that can be formed from the standpoint of regarding one's actions as up to oneself. (3) Agential knowledge is direct insofar as it is non-inferential, and self-warranted: it is not dependent on further beliefs to be in good epistemic standing. (4) Finally, that such beliefs yield genuine knowledge has been shown by appeal to a general virtue-theoretic account of epistemic warrant. In this way, the account also preserves what Setiya identifies as the central motivation for cognitivism, namely, the view that "the will is a capacity for practical knowledge" (2009: 131). For on the view I have defended, the will, like our perceptual faculties, is the source of a distinctive and irreducible kind of knowledge.

Chapter 6: Authoritative Knowledge

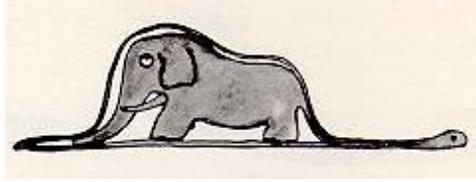
Philosophers working on practical knowledge tend to assume that such knowledge is restricted to what I have been calling ‘agential knowledge’, the knowledge that agents characteristically have of their own actions.¹³⁵ In this chapter, I challenge this assumption. I shall argue that there is a form of knowledge that agents have on the basis of their practical authority, a knowledge that need not even be about actions. And for all that, it is practical in Anscombe’s sense.

To introduce the notion, consider the following picture, which you may recognize:



If you do not recognize it, you might have good guesses as to what it represents: a hat, a slug, or a falling tent, perhaps. All of these guesses are, of course, incorrect. As we learn from the opening pages of *The Little Prince* (2015), the picture represents a boa constrictor eating an elephant. We are so told by the author of the picture, who draws a second one portraying the snake from the inside to enable us, unimaginative adults, to appreciate the true nature of his drawing:

¹³⁵ I do not believe the assumption was shared by Anscombe. As I noted in Chapter 1, one of her central examples of practical knowledge is the knowledge the director has of the nature of the construction she is guiding. This is an example of what I’m calling ‘authoritative knowledge’.



Suppose that you are one of the unimaginative adults in the story, that the author shows you the picture, and that after you say, ‘what a nice hat,’ he says, ‘it’s not a hat; it’s a boa eating an elephant’. The assertion is clearly true—but I want to claim something stronger: I want to say that this assertion expresses knowledge, a knowledge of a special type that the author has precisely in virtue of his authority as the author of the picture. I shall call this ‘authoritative knowledge’. The aim of this chapter is to present an account of authoritative knowledge.

I shall continue to work from within the Reidean tradition in epistemology. A central tenet of this tradition is that there is a set of privileged sources of information whose deliverances give us direct (non-inferential) knowledge of the world. The traditional list of sources includes perception, memory, testimony, reasoning, and reflection. In the previous chapter, I argued that we should include the will as a further knowledge source. The central thesis of this chapter is that authority is yet another basic knowledge source, one that also gives us practical knowledge.

If correct, the thesis is interesting in its own right. If authority yields knowledge in some way analogous to the direct way in which sources like perception and memory give us knowledge, that on its own would be an important epistemological finding. After all, the philosophical study of authority has predominantly fallen to ethicists and political philosophers, receiving little attention by epistemologists, unlike the other recognized basic

knowledge sources.¹³⁶ Moreover, I shall suggest below ways in which appeal to this notion might help to throw light on a number of important phenomena.

The paper will proceed as follows. I start by clarifying the nature and scope of the thesis that authority is a basic knowledge source (Section 1), and after distinguishing between an indirect and a direct model (Section 2), I offer two arguments in favour of the latter (Section 3). I then use the account of warrant from Chapter 4 to explain how it is that authority gives us direct knowledge in this way (Section 4). Finally, I close with some programmatic remarks on how we might employ the notion of authoritative knowledge to elucidate certain phenomena (Section 5).

1. AUTHORITATIVE KNOWLEDGE FURTHER EXPLAINED

The notion of authority that gives rise to authoritative knowledge is practical authority in general, of which artistic authority is only a species. To clarify the scope of the thesis that authority is a basic knowledge source, let me therefore present two more examples of authoritative knowledge that do not stem from artistic authority:

Captain: Captain O’Heguemon has command over the crew of a ship. Through a speakerphone in the lower cabin, she orders the sailors in the upper cabin to clean the floors of the main cabin before the day is over. Just on the basis of this order, and without performing any further checks, the captain comes to believe that the floors will be cleaned by the next morning. The crew is both competent and obedient, and they accomplish the task before the end of the day, as ordered. Here, the captain has authoritative knowledge that the floors of the cabin will be cleaned.

¹³⁶ It is important that I’m concerned with what is sometimes called ‘practical authority’, such as a captain has over her soldiers. This is to be distinguished from what is sometimes called ‘epistemic authority’, which is the power to speak and be trusted on a certain subject. A doctorate degree in biology, for instance, might give you epistemic authority, insofar as it entitles you to speak and be trusted about certain biological subjects, without conferring any practical authority, such as the authority to command others. See, however, Zagzebski’s (2015) pioneering work that appeals to the literature on practical authority to throw light on the notion of epistemic authority.

Single Father: A mother has died at childbirth before settling on a name for the child with the father. The father is now holding the baby, deep in thought with unnameable emotions, finding himself both a widow and a father. A nurse passes by and asks him, ‘What’s the name of the child?’. The father, who hasn’t given a thought to this until now, considers for a moment and replies, ‘her name is Juliana, like her mother’. This assertion expresses knowledge, a knowledge that he has in virtue of his authority over the name of the child. It is thus a case of authoritative knowledge.

We now have three examples of the target notion: the knowledge of the author of the picture from *The Little Prince*, the knowledge of the captain about the status of the floors in the morning, and the knowledge of the single father of the name of the child. These cases are simple: they are cases where the primary bearer of authority is a single person, and where the exercise of authority is successful. Cases of shared authority and failures raise complex problems of their own. Eventually I shall consider some of these problems (section 5), but in this chapter I am centrally concerned with the simpler cases. My aim is to develop a theory that fits these, that can *then* be used to assess the more complex issues raised by failure and shared authority. Finally, although there are obvious differences between the three examples, they should receive parallel epistemic treatment. Or so I shall argue.

Throughout the paper, I shall rely on an intuitive understanding of authority: authority is the power to settle a matter through one’s say. I take this understanding to be compatible with the central accounts of authority in the literature, such as Joseph Raz’s account of authority as the source of exclusionary reasons (1985a, 1981, 1985b, 2010), or Stephen Darwall’s theory of authority as the source of second-personal reasons (2011, 2010).¹³⁷ In determining whether someone possesses authority we can appeal to one of its

¹³⁷ Authoritative reasons have the form identified below as constitutive of second-personal reasons. When someone in authority exercises her authority, something ought to take place (and thus will, in the good case) in virtue of the fact that the person exercised her authority.

characteristic marks: a person in authority is one who is in a position to (legitimately) utter certain speech acts in imperatival form. ‘Let there be light!’, ‘Clean your room!’, or ‘Let her name be Juliana’, are all paradigm expressions of authority. When expressed by someone who lacks authority, such illocutions misfire.¹³⁸

Before defending the view that authority gives a special kind of knowledge, I need to defend the more basic claim that it yields knowledge. Although this might already be clear to the reader from the examples considered so far, let me give two further, and more principled reasons for this claim, focusing for simplicity on the case of the father.

First, consider a case where the father’s assertion starts a testimonial chain: after the father tells the nurse that the name of the child is ‘Juliana’, the nurse tells a doctor that that’s her name when he asks her, and the doctor believes her. It seems evident that the doctor thereby comes to know that the child’s name is ‘Juliana’, and that he comes to know this through testimony. However, it is widely held that for a listener to gain knowledge from a speaker’s testimony, the speaker must know what she asserts.¹³⁹ This means that the nurse must know that the name of the child is ‘Juliana’. Since she comes to know this also through a testimonial exchange, it seems that by parallel reasoning the father must know that that’s the name of his child. His knowledge, however, cannot be testimonial, nor does it seem to be gained from any of the standard knowledge sources. Rather, it seems

¹³⁸ The classical discussion of this issue is in Austin (1975: 28-29). A complicating factor that I shall ignore is that one can sometimes establish authority simply by acting as if one had it. See Thomason (1990), Witek (2013) and Langton (2015) for elucidative discussion of the issue.

¹³⁹ Though widely held (see e.g. Ross (1986); Burge (1993, 1997); McDowell (1994); Adler (2008)), the principle has been called into question by Lackey (1999; 2008, ch.3). However, the present example does not have the features that raise problems for the principle, so we can safely appeal to it in the argument.

clear that the knowledge is fundamentally grounded on his baptismal act, an exercise of authority.

The second argument appeals to the knowledge norm of assertion, employed in previous arguments. To recall:

Knowledge Norm of Assertion (KNA): a speaker can legitimately assert p only if she knows p .

When the father answers the nurse, he answers by asserting that her name is ‘Juliana’ (an assertion that also functions as an exercise of authority). Since this assertion seems legitimate, by KNA we can conclude that the father knows this is the name of the child; and, once again, he seems to possess this knowledge on the basis of his exercise of authority.

It is easy to see how parallel arguments could be given to establish that knowledge is what the child from *The Little Prince* and the captain have. But even if we grant that all of these are instances of knowledge, that still leaves open the central question I hope to answer in this paper, namely: What type of knowledge source is authority? Is it a source of indirect knowledge, merely giving us evidence on the basis of which an inference can be drawn, or a source of direct knowledge, so that authority is a basic knowledge source? I turn to this question next.

2. DIRECT AND INDIRECT MODELS

To decide whether authority gives us knowledge directly, we need a better understanding of the alternatives. The indirect model is an adaptation of Sarah Paul’s (2009a) inferential model for agential knowledge, explored in Chapter 3.

The direct and indirect models of authoritative knowledge differ with respect to what they take as the conditions that a person must meet for her belief to be warranted in a situation where she exercises authority. According to the direct model, a person wielding authority is epistemically warranted in forming a certain belief just on the basis of exercising authority. For instance, the captain is warranted in forming the belief that the main cabin will be cleaned just on the basis of ordering the sailors to clean it. The captain need not do any further epistemic work, such as reflecting on the fact that she is in a situation where her orders are reliably obeyed (so long as they are), nor does she need to check whether it has been obeyed before committing to the proposition. Of course, she might do all this, but the point is that to acquire authoritative knowledge, she is not required to do so. In this respect, exercises of authority are, on the direct model, on a par with exercises of our perceptual faculties as understood in the Reidean tradition: I am warranted in believing that there is a table in front of me just on the basis of seeing this table. I need not do any further epistemic work, such as checking whether my perceptual faculties are in working order, nor need I be aware of the properties in virtue of which perception gives us epistemic warrant (such as its reliability) in order to form a warranted belief.

The best alternative to the direct model—both for perceptual and authoritative knowledge—is an inferential model. On an inferential model, an exercise of authority is not on its own sufficient to provide epistemic warrant. Rather, such an exercise of authority serves at most as an evidential base for an inference. Thus, for Captain O’Heguemon to be warranted in forming the belief that the floors will be cleaned the next day, she would, at a minimum, have to reason in something like the following way: ‘I have ordered the sailors to clean the floors; the sailors generally follow the orders I give; therefore, the floors are

going to be cleaned”, and this conclusion will be based on the inference as a whole, rather than just on the exercise of authority.

According to the inferential model, then, to be warranted in forming a belief (and thus to obtain knowledge) on the basis of an exercise of authority, the person must go through a number of intermediary steps. At a minimum, the person must go through what I call an ‘inferential step’:

Inferential Step: If I exercise my authority in circumstances such as the present ones, *p* will obtain.

An inferential step is a belief that, through its content, appropriately links the exercise of authority with a state of affairs that is to obtain on its basis. The inferential step can occur unconsciously, without the person’s noticing it. However, since this is part of an inferential process, the inferential step must be occurrent; it cannot be merely dispositional, since the person must *use* this thought to arrive at the conclusion she reaches. This is constitutive of the notion of inferential warrant, which thus introduces two kinds of indirectness. First, there is mental indirectness, insofar as there is an intermediate mental state (the belief that constitutes the inferential step) between the exercise of authority, and the belief about how the world ends up being on its basis. Second, there is normative indirectness because the epistemic status of the belief depends on more than the exercise of authority: the belief can be in good epistemic standing only if such a step is taken.

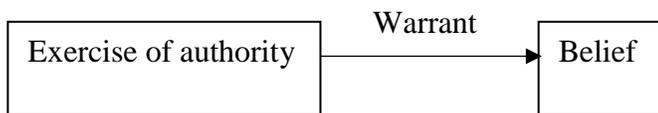
As I noted in Chapter 3, the inferential model imposes further requirements on warrant than merely thinking the inferential step. For instance, it also requires that the belief be drawn *on the basis* of the inferential step. Moreover, since an unwarranted belief cannot provide warrant for another belief, the conclusion reached on the basis of an inferential step won’t be warranted unless the inferential step is itself warranted too.

However, as different models might give different explanations of how these further conditions are met, my strategy will be to focus my objections to inferentialism on the more minimal condition identified above, namely, that a subject must think the inferential step to acquire warrant on the basis of an exercise of authority.

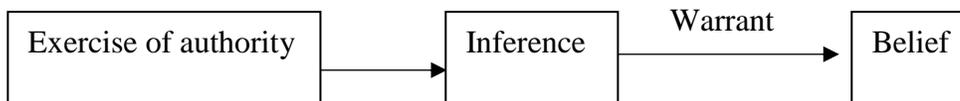
3. ARGUMENTS FOR THE DIRECT MODEL

The following diagram summarizes the differences between the two models, with arrows representing dependence relations between different states of a subject:

DIRECT MODEL:



INDIRECT MODEL:



There are many sources of information for which an indirect model seems correct. Consider, for instance, our desiderative faculty. Suppose that you are a heavy smoker, and you are craving a cigarette. Such a craving does not give you direct warrant to form the belief that you will smoke (if you formed the belief just on that basis it would be unwarranted); but under the circumstances, it might provide a good base for an inference which would yield warranted belief, and thus potentially knowledge that you will smoke. What about authority? Does it similarly require an inference to yield a warranted belief? I shall now provide two arguments against this view, by showing that exercises of authority can yield knowledge without the need for an inference.

3.1. Chatty Charlie's Assertion

The first argument is a slight modification of the one presented against inferentialism in Chapter 3, appealing to Chatty Charlie's assertions. To recall, Charlie is a person who cannot keep his thoughts to himself. All his thinking occurs publicly in speech form.

With this in mind, suppose you met Charlie at the hospital, and you knew his situation was that of the single father above: he just lost his wife at childbirth, and he is holding in his hands a yet unnamed child. You see him holding her and you ask him, 'What's her name?'. Charlie responds, 'He asked me what the name of the child is; well, we never settled this; what should I call her?' and, without saying anything else that is relevant to the case at hand, he replies after these musings, 'her name is Juliana, like her mother'.

How should we evaluate such an assertion? It seems clear that it is in good standing. After all, Charlie has full authority over the name of the child, so why shouldn't he be able to state what it is and settle it thereby? If the knowledge norm of assertion is correct, therefore, we can infer from the legitimacy of this assertion that Charlie possesses knowledge. This conclusion is in line with the direct account, according to which an exercise of authority is sufficient to provide epistemic warrant. Charlie's assertion merely constitutes a limit case, where the exercise of authority is self-justifying because it is both an exercise of authority and an assertion: qua exercise of authority, it provides epistemic warrant for the belief it expresses, ensuring that it can amount to knowledge.¹⁴⁰ Qua assertion, it is correct (by the lights of KNA) because it expresses or constitutes knowledge.

¹⁴⁰ Of course, since warrant is not *sufficient* for knowledge in my view, the belief amounts to knowledge also in virtue of other conditions obtaining (e.g. its not being defeated).

By contrast, an inferentialist account entails that the assertion is incorrect. This is because Charlie has not gone through an inferential step (we know this, because, *ex hypothesi*, he would have expressed the belief that constitutes such a step if he had), which means that he hasn't performed an inference, and therefore lacks warrant for what he asserts. As such, Charlie does not know, when he asserts that the child's name is 'Juliana' that that's her name. The assertion therefore falls afoul of the knowledge norm of assertion. Because it predicts, wrongly, that such an assertion is incorrect, we should reject the inferentialist account.

There are a number of ways of resisting this argument. I have tried to answer the most pressing ones in my discussion of the argument in Chapter 3. Hence, without further ado, I move on to the second argument for the direct model.

3.2 Stipulative Knowledge

The second argument for the direct model is based on one offered by Burge (1993) in response to Chisholm (1987) to show that memory gives us direct warrant. Chisholm argued as follows: Certain mathematical and logical proofs are extremely long and complex, requiring anyone running through them to employ her memory to keep track of all the steps. This employment of memory plays a central epistemic role: without it, the conclusions reached would be unwarranted. Hence, any knowledge acquired in this way must be supported by a premise to the effect that memory is a reliable source of information, such as:

(MEMORY): If I (seem to) remember that p , then p .

However, the warrant we have for a claim like (MEMORY) is at least partly empirical. Hence, Chisholm concludes that there are many mathematical truths for which “we cannot be said to have an a priori demonstration of the conclusions” (p.30), since the warrant we have for them is partly empirical.

Burge’s response flips this argument on its head. He argues that since it is evident that we have non-empirical knowledge of at least some of the conclusions that we reach as a result of long deductions, it must be that no such step as (MEMORY) is needed for us to be able to rely on the deliverances of our memory. Instead, he argues, memory is a content-preserving faculty that gives us direct access to the facts, and on the basis of which we are directly warranted in forming a belief without the need of further epistemic support.

A similar argument is available for thinking that authority gives us direct knowledge. To see this, consider stipulative steps, steps of the form “Let n be such and such”, which play a crucial justificatory role in many formal proofs. For instance, consider the following simple example of a proof showing that the area of an ABC isosceles triangle is 12 cm^2 , given that its base AC is 6cm long and its legs AB and BC are 5 cm long:

1. Let l be a line running perpendicular to AC that intersects B .
2. Let D be the point at which l intersects AC : evidently, D is the half-point of AC .
3. Let db be the segment of l from D to B .
4. Then, the length of $db=ABC$ ’s height.
5. DBC is a right angle, so by the Pythagorean Theorem, the length of $db=4$.
6. Hence, ABC ’s area= $4 \times 6 : 2 = 12 \text{cm}^2$.

The proof is incomplete, relying on intuition where we should expect explicit grounds (as in step 2), but it will do for our purposes. My interest is in steps 1-3, what I’m calling ‘stipulative steps’. These steps clearly play a crucial justificatory role in the proof. For instance, without step 3, step 4 would be unjustified, since db would lack a referent.

Stipulative steps seem to be clear instances of exercises of authority: by taking the stipulative step, the person performing the proof settles that certain terms refer to certain entities (such as a line segment). This explains why their characteristic expression: ‘Let db ...’ is in paradigm imperatival form (cf. ‘let them eat cake’, ‘let there be light’). Indeed, I am inclined to think that stipulative steps are just instances of baptismal acts, like the one performed by the single father in naming the child. The only difference is that in this case the baptized subject is an abstract entity rather than a person.

Suppose, then, that inferentialism about authoritative knowledge were true. This would mean that, to be warranted in believing on the basis of 3 that the line from D to B is db , one would have to think something like the following step, which would be part of the justification of step 3:

(DECISION): If I decide to name the line db , then its name will be db .

One problem with this is the same one that Burge identified in relation to memory: if (DECISION) is part of the justification of the above proof, the knowledge we would gain on its basis would seem to be empirical rather than a priori. However, there is an even more basic problem, both in this case and in the case of memory: the justification we gain by such a proof should not include *any* reference to mental states like memories or decisions, much less to the person performing the proof. The justification should consist entirely on the mathematical facts displayed in the proof. Since inferentialism implies otherwise, it should be rejected in favour of the direct model.

The foregoing argument seems to me decisive, both for memory and authority; but it is worth considering an objection that challenges its scope as concerns the latter. The objection is that the authority that we exercise when we reason on the basis of stipulative

steps is a very special one, which means that even if a direct model held true of it, we can't on that basis conclude that a direct model holds for all cases of authoritative knowledge. Thus, we might grant that a direct model holds for 'Stipulative Authority', the authority on the basis of which we reason with stipulative steps, but an inferentialist model might hold for the other sorts of authority we have considered.

My response is twofold. First, note that Stipulative Authority is not so different from other sorts of authority. In particular, it is natural to account for it in the same terms as the authority of the father; that is, as authority to name (a child/a mathematical entity). Second, anyone tempted by this line of response should consider whether they would draw the corresponding conclusion for the argument pertaining to memory: should we also say that there is a 'Proof-Specific Memory' which should be distinguished from ordinary memory as employed to remember e.g. recent moves in a game of chess, or the names of people we just met at a party? No: even if there are important differences between them—arising, for instance, from the nature of the particular objects remembered—the epistemic account should be the same for all these cases. Of course, this is consistent with thinking that there are important *psychological* differences between different kinds of memory (short-term/long-term; episodic/declarative; etc.). The claim is rather that insofar as they belong to the same genus, memory, they do so in part because they provide us knowledge directly.

The same holds for authority: the things over which authority is exercised in all the cases we have been considering have varied natures (they include the representational content of a picture, the name of a person, the state of a cabin, and the denotation of a term in a proof). As a result, there will be peculiarities to all these cases. In particular, there will

be peculiarities with respect to the conditions that must obtain for the exercise of authority to be successful, and hence for the beliefs formed on the basis of them to be true. However, as far as acquiring knowledge is concerned, they all seem just the same: an exercise of authority grounds a knowledge state by ensuring in a distinctive way that its content obtains. Hence, I am inclined to generalize from the conclusion reached on the basis of application of authority in mathematical reasoning, to its application in other domains.

Still, one might accept the claim that the cases of stipulative knowledge and the knowledge of the name of the child are analogous, but disagree about whether these should receive the same epistemic treatment as the case of the captain, for instance. For, one might hold, the cases of stipulative and baptismal authority are ones where a certain incorrigibility accompanies the use of authority: the exercise of authority, in these cases, entails that the relevant fact obtains, e.g. that *db* stands for the relevant line, or that the name of the child is 'Juliana'.

However, I dispute whether exercises of authority in these cases ensure incorrigibility. Imagine that as soon as the father says, 'her name is 'Juliana' (and forms the corresponding belief) he suffers a stroke. The nurse mishears him saying, 'her name is 'Juana'', and as a result, the wrong name is placed in the birth certificate of the child. In that case, the father's belief, formed on the basis of authority, would be false.¹⁴¹ Or again, suppose that the person exercises authority, but it is disregarded (as in the case of the baptism of John the Baptist that I discuss in the last section). Once again, forming a belief on the basis of exercising authority would yield a false belief.

¹⁴¹ Another way in which one might form a false belief is if one mistakenly uses one's authority. For instance, the person performing the proof might use a variable that she is not permitted to use at that step of the proof (e.g. because it already has a determinate referent).

More importantly, it is implausible to think that incorrigibility would give us enough grounds to differentiate between epistemic sources. Take the thoughts, *I am thinking a thought* and *I am thinking about a tiger*. Presumably, the capacity by which one knows that one is thinking these thoughts is the same. However, the former thought is incorrigible, while the latter is not: if one thinks, ‘I am thinking a thought’ the thought is necessarily true, since it is self-verifying; but if one thinks ‘I am thinking about a tiger’ one could be mistaken by thinking instead about a reptile that looks like a tiger (Putnam 1975). Hence, even if we accepted that thoughts formed on the basis of an exercise of stipulative authority are incorrigible and thoughts formed on the basis of an exercise of a different kind of authority are not, this would not show that they should receive different epistemic treatment. Further support for a unifying treatment will be given in the next section, by providing a unified explanation of how exercises of authority as a whole can be the proper grounds for non-inferential knowledge.

4. DIRECT KNOWLEDGE IN THREE VIRTUOUS STEPS

The two foregoing arguments seem to me to present a strong case for a direct model of authoritative knowledge. However, they do not yet give us much insight into the features of authority in virtue of which it is capable of giving us such direct warrant. In this section, I attempt to provide an account that yields such insight by appeal to the account of warrant developed in Chapter 4.

To recall, I defended the following account of epistemic warrant:

Warrant Through Competence (WTC): *S* is warranted in forming/maintaining the belief that *p* iff (a) *S*’s belief is formed/maintained through a performance that is reliable at ensuring that *p* is true, and (b) it is so reliable in virtue of being guided by the aim of ensuring that *p* is true.

To evaluate whether one is warranted in forming a belief on the basis of an exercise of authority, let us focus on the case of captain O'Heguemon. Like perceptual beliefs, beliefs formed on the basis of authority seem to have the following tri-partite structure:

4. *Information-check*: A system ensures that certain information holds true of the world.
5. *Delivery*: The information is delivered in an appropriate form to be used at a next stage of the process.
6. *Dependent Task*: A performance of a determinate sort is taken on the basis of the information thus delivered.

In the case of the captain, the information-checking system is the system of authority in virtue of which the captain has command over the sailors. This system aims to ensure that certain information holds true of the world (whatever the captain orders within her domain of authority). Hence, when the captain gives an order, this constitutes the beginning of a joint performance that (if things go well) culminates in the joint action of executing the order. Like perceiving, giving an order is generally a conscious performance. Consciously keeping track of what one orders ensures that the relevant information is available for the next stage of the performance, doxastic consideration. In this third stage, the captain is in a position to form a belief on the basis of her order. In the case at hand, she is in a position to form the belief that the cabin will be cleaned by the next morning.

Finally, these steps together constitute a performance part of the function of which is to enable the captain to form true beliefs about matters over which she holds authority. Unless she is able to form such beliefs, it is hard to see how she could exercise her authority appropriately. For instance, suppose it is important that the floors get cleaned by the next morning. If the captain could not form the true belief (indeed, know) that it was going to be cleaned by the sailors, she should arrange for it to be done otherwise. This is not an

isolated example. Unless authorities are generally in a position to know that their orders will be obeyed, it is hard to see how they could exercise authority appropriately.

The importance of this point comes out in a case discussed by Louise Antony and Rebecca Roman Hanrahan (2005): A woman professor has prepared an exercise for her class that requires her students to form a circle. Hence, at the start of the class, she tells the students to form a circle. But the students do not do this. Instead, they proceed to challenge the teacher's authority: *Why should they form a circle? What would that add to the class?* Hanrahan and Antony argue that women teachers are particularly prone to experience these situations.¹⁴² What I want to highlight here is just how much harder it is for teachers who cannot presuppose that their authority will be respected to plan a class. A male teacher whose authority is respected can simply take for granted that his class will form a circle, and plan the class accordingly. By contrast the female teacher whose authority is disrespected must plan for many uncertainties: will her students start challenging her? What are these challenges? Should she attempt to address them, or ignore them? If she addresses them, how should she do so? Etc. If she is teaching a class on authority or a related matter, addressing these questions might be a helpful exercise. But quite generally the inability to know that her orders will be obeyed constitutes an obstacle to carrying out a good class, which is the fundamental reason why the teacher possesses such authority to begin with. Authority is tasked with effective planning; and effective planning requires knowledge that authority will be respected; hence authority requires knowledge that authority will be respected.

¹⁴² A recent study that showed that women were more prone than men to be subject to unreasonable demands from their students supports this view. See El-Alayli, Hansen-Brown, and Ceynar (2018).

What remains to be shown is that the tripartite performance of the captain is competent relative to its epistemic goal. As with the case of perception, I shall appeal to the following principle:

Competence by Parts (CP): Given a complex performance p with an overall aim A that is constituted by a series of sub-tasks $\langle t_1, t_2, \dots, t_n \rangle$, p is competently performed (by the standards of A) iff the subtasks in the series, $\langle t_1, t_2, \dots, t_n \rangle$, are competently performed.

First, then, given that the captain is in a position of authority and that the sailors are competent and obedient, her order ensures that the cleaning of the cabin gets carried out competently and obediently. Thus, the first task of information-checking is done competently. Of course, what was true of the case of agential knowledge is true here: the way in which the information is checked is practical rather than theoretical, as it was in the perceptual case. Whereas when we exercise our perceptual systems, these adjust so as to match the world, when we exercise our authority the world adjusts to match our dictates. This is part of what makes the knowledge available thereby practical. Second, given that generally a person is conscious of the orders that she gives, when the captain gives the order (e.g. 'clean the main cabin!'), the information that the cabin will be cleaned is made available for doxastic consideration. Finally, forming beliefs on the basis of such conscious awareness of orders reliably yields true beliefs in the present circumstances. And the reliability will be fundamentally explained by the fact that the tri-partite performance as a whole is designed to adjust to ensure that such beliefs are true. For instance, if the sailors encounter difficulties in accomplishing the orders, they will attempt to find ways to overcome them, until the orders are fulfilled. Hence, the resulting beliefs are competently formed, and hence directly warranted on the basis of exercising authority.

5. IMPLICATIONS AND COMPLICATIONS

I began by introducing the notion of authoritative knowledge, the knowledge that the person in authority has on the basis of exercising her authority, with a series of examples. I then presented an indirect, inferentialist model, and a direct model as rival views to account for this sort of knowledge, and I argued that the direct model is to be preferred. Finally, I provided a virtue theoretic account that showed which features of authoritative exercises enabled them to provide direct warrant.

As mentioned at the outset, I am interested in the phenomenon of authoritative knowledge partly because of its connection to the notion of agential knowledge that has been central in the literature on intentional action since Anscombe's seminal *Intention*. In brief, the relation between the two seems to me to be the following. Both are practical sources of knowledge, insofar as the warrant involved has a world-to-mind rather than a mind-to-world direction of fit, in the way just explained at the end of the last section: exercises of authority and intentions provide warrant for belief because they competently ensure that certain information holds true by making it the case that it does. However, whereas the scope of agential knowledge is restricted to knowledge of one's actions, the scope of authoritative knowledge can extend to the actions of others, and even to non-agential facts (like the fact that a person's name is 'Juliana' or that a term denotes a certain line segment).

The notion of authoritative knowledge may also help to throw light on other phenomena, like semantic knowledge. It has been thought that one is always in a position to know what one means: meaning is luminous. This view gets its purchase from the fact that what we mean in a context of conversation is quite generally up to us. Just as the author

of the image with which we began can decide what the image represents, we can generally let our words mean what we want in a conversation. As such, the knowledge that one has of what one means is (at least partly) authoritative. However, the qualification (“generally”) is important: language is a communal phenomenon, so no particular speaker has sole authority over what is meant by certain signs—she shares that authority with other speakers of the language. As such, one cannot always assume that one means what one means to say, contrary to the thesis that meaning is luminous. Recognizing semantic knowledge as a species of authoritative knowledge allows us to account for what is special about semantic knowledge without holding an implausible luminous thesis.¹⁴³

The case of linguistic meaning highlights only one of the complications that may arise when authority is shared. Another complication arises from the fact that to be effective, authority must be recognized, which it often isn’t as in the case of the baptism of John the Baptist, as told in the New Testament:

And it came to pass, that on the eighth day they came to circumcise the child; and they called him Zacharias, after the name of his father. And his mother answered and said, ‘Not so’; but he shall be called John. And they said unto her, ‘There is none of thy kindred that is called by this name’. And they made signs to his father, how he would have him called. And he asked for a writing table, and wrote, saying, His name is John. And they marvelled all. (Luke I:59-63)

In this case, Elizabeth attempts to name her child ‘John’. However, her authority to name the child is unjustly unrecognized by her audience. As such, they look for an authoritative source they recognize, Zacharias, the father. Only once he exercises his baptismal authority is John’s name recognized as such. In this case, there is an unjust distribution of epistemic

¹⁴³ Similarly, I think the account can help to explain how an author has a distinctive knowledge of her own works, while granting that this knowledge is restricted, for instance, by the authority of the audience. I hope to explore this idea in future work.

resources: Zacharias's authority is a source of direct knowledge, whereas Elizabeth's is not, because his say and not her say settles the name of the child.¹⁴⁴

I am inclined to think of these cases as deviations from the standard one, where authority is obeyed, and the person can plan accordingly. As such, I am inclined to treat them, as far as the epistemology goes, as the analogues of illusions and hallucinations in the perceptual case. This is a point that bears further discussion. However, since this is also a good place to stop, I present one last illustration of the phenomenon I have aimed to elucidate: I can tell you—because I know so—that the paper ends here.

¹⁴⁴ I used to think this would count as an example of epistemic injustice in Fricker's sense (2007). However, because the primary way in which the person is harmed in this way is not by means of an epistemic exchange, I no longer hold this view.

Chapter 7: Freedom and Assertion

In recent years, several philosophers have defended what is known as an ‘interpersonal’, ‘Strawsonian’, or ‘assurance’ view of testimonial knowledge.¹⁴⁵ According to this view, testimonial knowledge is distinctive insofar as its acquisition requires the speaker and listener to engage in what I shall call, following Stephen Darwall (2006), a ‘second personal’ transaction, such as obtains when someone makes a promise. Interpersonal views are often motivated by appeal to cases where it seems rational to believe against the evidence (for instance, where a friend asks us to believe that he will quit smoking), as such cases suggest that testimonial transactions are governed by interpersonal rather than evidential norms.¹⁴⁶ However, opponents of the view are understandably skeptical of the force of such arguments: some doubt that it is rational to believe in this way; others wonder whether, even if it is rational, it is *epistemically* rational to do so, since these might be cases where practical reasons outweigh epistemic reasons.

In this chapter, I present a new argument for an interpersonal account, one that relies on considerations about essential properties of testimonial exchanges rather than on intuitions about cases where we seem to judge against the evidence. Central to the argument is the contention that there is a tension between two of the distinctive features of testimonial

¹⁴⁵ Recent defenders of interpersonal views include Watson (2004), Hinchman (2005, 2014), Moran (2005, 2013), Faulkner (2007, 2011), McMyler (2011, 2013), Marušić (2015, 2017), and Baker and Clark (2017), but some of the central ideas associated with interpersonal accounts are already present in work by Anscombe (1979), Baier (1986), Ross (1986), and Baker (1987). Holton (1994) is sometimes included as a defender of interpersonal views, but since he is a non-cognitivist about trust, I don’t count him as a defender of such a view about testimonial *knowledge*. Of course, there are important differences between the interpersonal views defended by these writers.

¹⁴⁶ See e.g. Baker (1987) and Marušić (2015, 2017), though the full force of the arguments depends not only on the intuitions about the cases, but also on the framework offered to account for them.

knowledge. First, testimonial exchanges are supposed to give the listener direct (i.e. non-inferential) epistemic warrant, based only on another's say-so.¹⁴⁷ Call this the *Directness Requirement*. Second, testimonial knowledge is supposed to be gained on the basis of a personal engagement. Call this the *Personal Requirement*. The tension arises from the fact that, to meet the directness requirement, we must regard the speaker's assertion as having a determinate aim, namely, of providing us with true beliefs, but to meet the personal requirement we must regard the speaker as a free being—free, in particular to set her own aims for her actions. Since assertions are actions, it would seem that to engage in a testimonial transaction we would have to regard the speaker as both having a determinate aim for what she asserts, and as having no such particular aim, since assertions can be used for any number of other purposes among which the speaker, as a person, is free to choose.

I argue that to resolve the tension, we must appeal to *tellings*, understood as intrinsically second-personal assertions. When a person tells another something, they bind themselves to a certain standard as their governing aim: the standard of passing along true information, such that the other will be able to form a true belief just on the basis of the assertion. Because a telling is freely undertaken, one can regard a speaker who tells one something as a person, while at the same time regarding her as bound by an epistemic norm, enabling one to form a belief just on the basis of her say-so. An important advantage of my view is that it shows how second-personal reasons can play an epistemic role in the

¹⁴⁷ The claim that another's say so gives us non-inferential warrant is rejected by reductionists about testimony, who will require that the speaker make an inference using the assertion as a base. I don't aim to defend anti-reductionism in this paper: I am convinced of this view by the arguments in Coady (1992), so I shall be taking the view for granted. Nevertheless, the central premises of my argument are ones that even anti-reductionists could endorse, so it is meant to have wide appeal.

acquisition of testimonial knowledge. It thus enables us to meet a powerful objection raised by Jennifer Lackey (2008) against interpersonal views.

I'll proceed as follows. In Section 1, I explain interpersonal accounts in more detail, and consider some of the reasons that have been given for holding such accounts. In Section 2, I summarize Lackey's challenge, and argue that the responses that have been offered to it are unconvincing because they fail to show how a second-personal factor could be of *epistemic* significance. In Section 3, I argue that to offer direct warrant, a deliverance of information must aim at providing the speaker with true beliefs. In Section 4, I argue that when the process is a communicative one between two persons, it requires speaker and hearer to engage in a second-personal interaction, because this is the only way to regard another as a person while also regarding her as having the aim of delivering true information. In Section 5, I briefly explore how the view defended here bears on the question of whether epistemic partiality is rational.

1. INTERPERSONAL VIEWS AND THEIR PLACE IN THE LITERATURE

Accounts of testimonial knowledge in the literature are usefully divided into two camps. According to the first, *reductionist* camp, whose roots are often traced back to David Hume, there is no distinctively testimonial knowledge: true, we can often gain knowledge on the basis of what others say, but this knowledge is of the standard inferential variety.¹⁴⁸ As such, we are justified in believing what others say only to the extent that we have independent evidence to think that what they say is likely to be true. According to the

¹⁴⁸ For Hume's views on testimony, see especially *Treatise* 1.3.4 SB 83-4 (2007), and especially essay X 'Of Miracles' in the *Enquiry* (1904). A sustained contemporary defense of reductionism is found in a series of articles by Elizabeth Fricker (1987; 1994, 1995, 2004, 2006a, 2006b).

second, *non-reductionist* camp, whose roots are often traced to Thomas Reid, testimony is a basic knowledge source, and its acquisition cannot be explained just by appeal to other types of epistemic sources (it is, to this extent, irreducible).¹⁴⁹ Crucially, testimonial knowledge on this view is directly warranted on what another says, rather than inferentially based.

This characterization of the views is deliberately vague, and importantly non-binary, since there are mixed accounts of testimonial knowledge in the literature with features from both sides.¹⁵⁰ Be that as it may, I shall take interpersonal views to fall squarely into the non-reductionist camp, as they include a commitment to the following claims:¹⁵¹

- (1) *Distinctiveness*: the epistemic features of testimonial belief cannot be fully explained by appeal to epistemic sources of other types.
- (2) *Directness/non-inferentialism*: the kind of justification that testimony gives us is direct (non-inferential): a hearer is (at least sometimes) warranted in forming a belief just on the basis of someone's say-so.
- (3) *Non-evidentialism*: The norms that govern the formation of beliefs in testimonial transaction are of a distinct, non-evidentialist, sort.

Indeed, we can think of interpersonal views as providing an explanation for why (1)-(3) hold, an explanation that gives these ingredients a positive characterization. That

¹⁴⁹ For Reid's views on testimony, see especially *IHM* VI.xxiv (2000) . A sustained contemporary defence of anti-reductionism inspired by Reid is offered by Coady (1992). See also Burge (1993, 1997), and McDowell (1994), among many others.

¹⁵⁰ For instance, Lackey (2008) and Goldberg (2006) explicitly present their accounts as mixed, and Sosa's (2006) account would also count as mixed by this characterization.

¹⁵¹ However, note that Fricker (2006a) offers an interpersonal account that does away with Directness, but all other defenders of interpersonal accounts, as far as I know, hold this requirement.

testimonial knowledge is distinctive, direct, and non-evidentialist is to be explained by the fact that its acquisition is grounded in a second-personal transaction.

To elucidate what this means, I will first offer a characterization of second-personal transactions generally, and then explain how it applies to the case of testimony specifically. Following Darwall (2006), I take second-personal transactions to be defined in terms of a special kind of reasons, second-personal reasons. On Darwall's formulation, "[a] second-personal reason is one whose validity depends on presupposed authority and accountability relations between persons and, therefore, on the possibility of the reason's being addressed person-to-person" (2006: 8). The key notions here are those of authority, accountability, and person-to-person address. It is helpful to explain what this means using a particular example. Suppose Ann promises Bob that she will Φ , and Bob accepts the promise. To be successful, such a transaction requires Ann to address Bob as a person capable of promise uptake; and, correspondingly, it requires that Bob regard Ann as a person, capable of making promises. Moreover, to understand the specific transaction that takes place we must appeal to the notions of authority and accountability: if the transaction succeeds, Ann thereby becomes accountable to Bob to Φ , and Bob has authority to demand of Ann that she Φ . Finally, if the promise is accepted, Ann is thereby *bound* by this promise: whereas, all else equal, she would have been free not to Φ ahead of making the promise, if she fails to Φ after the promissory transaction occurs, she would fail by the norm that she has placed herself under by her very act of promising, wronging Bob.

My preferred definition of second-personal transactions proceeds by identifying a distinctive *form* of reason-giving, one that becomes apparent when we contrast two ways

in which a person can give another a reason.¹⁵² On the one hand, one can give someone a reason to do something by simply pointing out a certain consideration that bears on the person's plans and interests. For instance, one might give someone a reason to call a cab by pointing out that the next train leaves in 50 minutes, or a reason not to cross a street by pointing out that he might get run over by an approaching truck. On the other hand, one can give someone a reason to do something by exercising one's authority to do so, a reason that will thus be grounded in that very exercise of authority. For instance, a mother can give her daughter a reason to call a cab by requesting her to be home for dinner, and a policeman can give someone a reason not to cross the street by ordering him not to do so.¹⁵³ In the first kind of case, one gives a reason by directing the other's attention *outwards*, as it were, a reason that is grounded on facts external to the very act by which one makes those reasons available. By contrast, in the second way one directs the other's attention *inwards*, as it were, insofar as the reason offered is grounded in the very act by which it is made available. Second-personal reasons, then, are those for which the very act that makes the reasons available is itself a fundamental ground of the reasons thus made available. This is why if the daughter asked why she should come home, the mother should say,

¹⁵² Though I am here departing from Darwall's treatment in *The Second Person Standpoint*, the present characterization of second-personal interactions is consistent with his discussion in the book, and has been influenced by the contributions by Thompson (2004), and especially Darwall's "Because I Want It" (2001). In this essay, Darwall comes close to the characterization I give in terms of a form of reasons-giving, though, if I understand it, the view presented is one where the reasons simply have a characteristic *content*, one that refers back to a person's desires (instead of the content of that desire). I suspect this formulation loses importance in the book-length treatment of the topic precisely because a characterization in terms of content alone does not identify what's distinctive of second-personal reasons.

¹⁵³ Notice that this form of reason-giving is common to two species of second-personal reasons that Darwall distinguishes, namely, 'juridical' and 'of the heart' (Darwall 2017). Juridical reasons are in the form of commands, but requests, I take it, would be 'of the heart', because the person requesting something generally won't be in a position to demand it. Still, when I request you to Φ , I give you a distinctive reason to Φ that is grounded in the very request.

‘because I’m asking you to do so’, and if the person asked why he shouldn’t cross the street, the policemen should say, ‘because I’m ordering you not to do so’. By contrast, the fundamental ground for the reasons made available in the non-second personal cases are only the facts that are pointed out: the fact that the next train leaves in 50 minutes, or that the person will get run over if he stays on the street. The act by which these reasons are made available is not a fundamental ground of those reasons.¹⁵⁴

The above should suffice as a characterization of second-personal interactions for the purposes of this paper. Given this framework, the question I am interested in is: are the reasons that ground testimonial knowledge genuinely second-personal? One reason to think so is that these interactions seem, on the face of it, to display the form identified above as distinctive of second-personal exchanges. Thus, suppose your friend Francisca has told you that *p*, and you form the belief that *p* on that basis. Now suppose someone asks you why you believe *p*. It seems natural for you to reply that you do so *because Francisca told you*. This mirrors the structure that we find in interpersonal interactions generally, offering a certain form of engagement as the fundamental basis for doing something. Moreover, as is characteristic of these engagements, when someone forms a belief on the basis of one’s testimony, other forms of second-personal interaction become appropriate: for instance, if the person has lied, or has been negligent in the way in which they acquired the information, one is entitled to hold them accountable with blame.¹⁵⁵

¹⁵⁴ Of course, there are cases where the act itself will be a ground accidentally, because it is part of the grounds made available: e.g. if someone is wondering if anyone in the room can speak, I can give her a reason to believe this by saying ‘I am speaking’. The assertion itself verifies the asserted content. These are inward looking *by accident* (*kata sumbebekos*) in the Aristotelian sense in which the patient and the doctor might be identical.

¹⁵⁵ Relatedly, as Anscombe noted (1979), it is sometimes disrespectful to distrust someone who offers testimony (as it is sometimes disrespectful to reject a promise).

The second-personal structure of testimonial engagements can in turn be used to explain the distinctiveness, directness, and non-evidentiality that non-reductionists hold are essential to testimonial knowledge: testimonial knowledge is distinctive insofar it gives us a second-personal basis for belief; it is direct, insofar as the person who forms the belief can simply rely on another's say-so to permissibly form her belief, rather than on independently available evidence; and it explains why testimonial exchanges are not governed by evidential norms, since second-personal exchanges in general are not governed by such norms, but by distinctively second-personal ones.

It has also been thought that second-personal accounts can provide an attractive solution to a puzzle that arises in cases where one person tells another something that seems to conflict with a lot of independently available evidence, as in the following cases:¹⁵⁶

Accused Friend: Your friend has been accused of a serious crime, and the evidence overwhelmingly suggests that he is guilty. So much so, that you think to yourself: 'If I just considered the evidence on its own, I would think he's guilty'. However, your friend has told you that he is innocent, and, on that basis, you believe that he is.

Quitting Addict: You know the statistics: people trying to quit smoking are overwhelmingly likely to fail. However, after talking about it for a while, your spouse wakes up one day and tells you, "I'm quitting smoking". On that basis, you believe that she will.

These cases are interesting because they seem to be ones where agents justifiably form beliefs that are contrary to the evidence. The interpersonal view of testimonial knowledge seems to offer an attractive explanation for this: since these are second-personal engagements, they are ruled not by evidential norms, but by second-personal norms. As

¹⁵⁶ See especially Baker (1987) and Marušić (2012, 2015) for appeals to such cases to motivate interpersonal views. On epistemic partiality more generally, see also Keller (2004), Stroud (2006), and Paul and Morton (2018).

such, believing against the evidence does not violate any norms, so long as the belief is formed in accordance with second-personal norms.

Finally, interpersonal views seem well-suited to explain two perceived contrasts. First, there seems to be an important difference between the way in which we acquire knowledge from other people and the way we acquire knowledge from mere instruments, like thermometers.¹⁵⁷ Although we can acquire knowledge of the weather just by looking at the thermometer, or by being told, these ways of acquiring knowledge seem importantly different. On the interpersonal view, the difference is that only the latter requires you to engage in a second-personal transaction. Second, there seems to be an important difference between the way we acquire knowledge from being *told* something, and the way we acquire knowledge when we merely *overhear* someone.¹⁵⁸ If you overhear your friend at a bar telling your other friend something, it seems like you and your other friend stand in a different relation to what was said: what they say seems to be *mere* evidence to you, whereas your friend seems to have a distinctively testimonial reason to form the belief. On the interpersonal view, the difference is that only the friend who is told acquires knowledge by engaging in a second-personal interaction, rather than by merely coming upon evidence that justifies the belief. Hence, only the friend has testimonial knowledge. Arguably, the two contrasts are closely connected: only when you form a belief just on the basis of being

¹⁵⁷ Anscombe (1979), Ross (1986), Moran (2005, 2013), and Hinchman (2005, 2014) all emphasize the differences between these two cases, though see Sosa (2006) for a differing view. Sosa's account is criticized by Wright (2014b), persuasively in my view.

¹⁵⁸ See Hinchman (2005, 2014), Moran (2005, 2013), and Faulkner (2007) for views that emphasize this difference, and Lackey (2008) and Weiner (ms) for criticism. I do not think the issue can be resolved just by appeal to intuitions about cases, and my grounds for siding with those who recognize an important difference will be provided when my account of testimonial knowledge is on the table, in section 5.

told do you treat the speaker as a person rather than as an instrument for the acquisition of knowledge.

However, none of these arguments are decisive, and there are obvious objections to each of them. Working backwards, some doubt whether there is a disanalogy of epistemic significance between being told something and merely overhearing it, or between knowledge acquired from instruments, and knowledge acquired from persons. Similarly, some would argue that in cases of epistemic partiality, like **Accused Friend** and **Quitting Addict** you should not believe the relevant persons; and among those who think that you should believe them, many would hold that this is only because these are cases where other norms—such as norms of friendship—trump epistemic norms. Non-reductionists who reject interpersonal accounts might offer alternative explanations for *Directness*, *Distinctiveness*, and *Non-evidentialism*, and reductionists would of course reject the need to account for them. Finally, one could hold that even if tellings are second-personal in nature and have important connections to other interpersonal attitudes, this does not show that these features play a role in the acquisition of knowledge.

This last is, in my view, the gravest problem: none of these arguments offer insight into how second-personal features could play a specifically epistemic function. My view is that the arguments cannot on their own provide support for interpersonal views until such an account is given. Moreover, while I take most of the above as advantages of second-personal views, I remain ambivalent about others, such as how to think about cases of epistemic partiality, a question I shall return to at the end of the paper. Before presenting my positive account, however, I shall first consider an argument by Jennifer Lackey (2008, ch.8) that aims to show that second-personal features cannot play an epistemic role.

2. LACKEY'S DILEMMA

Lackey's objection is in the form of a dilemma, as follows. The epistemic warrant that a listener gets from being told that p either requires her to meet non-second-personal conditions for warrant, like reliability and good evidential support, or it does not. *Horn 1*: if the non-second-personal conditions must be met, then they seem to do all the epistemic work, and second-personal reasons do not add anything of epistemic significance. *Horn 2*: if the non-second-personal reasons need not be met, it is hard to see how a purely second-personal reason, such as the granting of a distinctive assurance from the speaker, could add anything of epistemic significance. *Conclusion*: interpersonal factors play no epistemic role in the acquisition of testimonial knowledge.

Expanding on the set-up, the argument does not presuppose any particular account of epistemic conditions. These could include externalist conditions like reliability or causal dependency, or internalist conditions like Cartesian justification or anti-defeat. The argument does presuppose, however, that we can identify these conditions, and that second-personal factors are not required for them to obtain. To anticipate, it is this last assumption that I shall be calling into question later.

Grant, then, that as defenders of interpersonal views hold, testimonial exchanges are constituted by second-personal interactions. The dilemma confronts us when we ask whether the second-personal factors make an epistemic contribution. *Horn 1*: Suppose that in the cases where the transaction well (when the hearer acquires testimonial knowledge), the properly epistemic conditions must be met. When this is the case, these properly epistemic conditions seem to fully explain why the listener acquired knowledge. If the

agent's belief is true, appropriately justified, reliable, evidentially supported, etc., what else could be needed for it to constitute knowledge? Even if there is a second-personal aspect to testimonial exchanges, this would not add anything of epistemic value. That a second-personal reason is offered to the listener might explain why we are entitled to blame others when they lie to us, or why a speaker might feel let down when she is not believed; but the reason is not epistemic since it plays no role in the explanation of why the hearer's belief is in good standing. The explanation for the latter is fully in terms of traditional epistemic properties.

Horn 2: Suppose, on the other hand, that the standard, properly epistemic conditions that must obtain for one's belief to be in good standing need not obtain in cases of testimonial belief. Then it is hard to see how the belief could be in good epistemic standing (even if in good standing relative to other standards, like friendship, or even, rationality *tout court*). As Lackey notes, on the assumption that the second-personal element enters the picture when a speaker invites a listener to trust her:

[The speaker's] inviting [a listener] but not [an overhearer] to trust him does not make it more likely that the testimonial belief in question is true for [the listener] but not for [the overhearer]—they are both receiving testimony with the same degree of reliability, the same kind of truth-tracking, the same amount of proper-functioning, and so on. (2008: 14)

Although Lackey is here arguing against the specific contention that second-personal reasons are given only by one who *tells* another something, the conclusion is more general: Second-personal features, as such, do not seem to add anything of epistemic value. Hence, if we want to discern what is *epistemically* distinctive about testimonial knowledge, second-personal features “drop out of the picture” (p.238).

Responses have been offered to this problem, but all of them seem to me unpersuasive.¹⁵⁹ One strategy is to hold that in testimonial interactions the norms shift, so that we should form beliefs on the basis of second-personal (instead of e.g. evidential) reasons.¹⁶⁰ However, this strategy suffers from a now familiar problem: even if the norms differ, it is not clear that someone who forms beliefs on the basis of those norms behaves epistemically correctly. A second response appeals to assurance: when we offer testimony, the listener can hold us accountable if we lie. However, if assurance is understood as a mere outcome of the procedure,¹⁶¹ the assurance works merely as a further piece of evidence for the listener: one is warranted in forming the belief because one knows the speaker has a good reason not to lie; but this is not epistemically different from believing someone because one knows she is under threat if she lies. Finally, what I regard as the most promising response appeals to the notion of ‘epistemic buck-passing’: by engaging in a second-personal transaction, the listener is able to justify her testimonially formed belief on the basis not of grounds available to her, but the speaker’s grounds.¹⁶² The main problem with this response is that it isn’t clear that a second-personal interaction is necessary to pass the buck in this way. On the view under consideration, when I form a

¹⁵⁹ The objections I present are given in only rudimentary form, but I hope in a way that makes their force clear.

¹⁶⁰ A version of this response is given by Marušić (2015).

¹⁶¹ The qualification is needed because on the account I offer, assurance plays a distinctive role, but not as a mere result, but as shaping the form of the transaction in such a way that it can provide direct warrant. Moran (2005, 2013) is sometimes interpreted as holding the external result view; but I think he is better understood as pursuing the first strategy according to which non-evidential norms govern testimonial exchanges.

¹⁶² Epistemic buck-passing was initially defended by Goldberg (2006) as a way for reductionists to recognize the distinctiveness of testimony, but McMyler (2013) and Baker and Clark (2017) argue that the kind of buck-passing distinctive of testimony can only be acquired on the basis of an interpersonal exchange. My thinking about this strategy and about testimony in general is indebted to discussions with Phil Clark.

belief that p on the basis of S 's telling me that p , and my belief is called into question, I do not have to be able to defend my belief myself. Instead, I can defer to S for the justification. In such circumstances, it seems appropriate to justify one's beliefs by saying, 'that's what S said, and I assume she knows—if she doesn't, my belief is unjustified'. The problem is that we seem to be able to pass the buck in the same way when mere instruments are involved. For instance, asked to justify my belief that the temperature in the room is 64 degrees, I could point to the thermometer and say, 'that's what the thermometer reads, and I assume it's working properly—if it isn't, I suppose my belief is unjustified'.¹⁶³ We can also do this with people, even when there is no question of a second-personal interaction. For instance, suppose several people come running out of a theater, screaming in fear. On that basis, I would seem justified in thinking that there's danger inside. Asked to justify why I thought so, I might say, 'They seemed to think so! If they weren't justified in thinking it, I suppose my belief is unjustified.' Hence, second-personal transactions seem unnecessary to pass the epistemic buck.¹⁶⁴

¹⁶³ You might say: 'thermometers don't have mental states, so applying the notion of justification to them makes no sense' (cf. (Wright 2014b)). This response depends on an internalist account of justification, which I would reject; in any case, the next example is meant to address this issue.

¹⁶⁴ Another attractive account is defended by Hinchman (2014), who argues that testimonial transactions are distinctive because the speaker represents herself as giving the listener warrant to close doxastic deliberation (by forming a belief) in part by presenting the information in a caring way that takes into account that person's plans and interests. The argument is centrally motivated by the apparent fact that in some contexts, a speaker should not tell a listener what she is entitled to assert in other contexts: e.g. if Ann asks Bob if the weather is mild in NYC at this time of the year, and Bob knows it is (and tells this to other people), but also knows that Ann is particularly sensitive to cold weather and is trying to decide whether to bring a jacket, Bob should not say that the weather is warm. Hinchman argues that in giving the answer that is appropriate to the hearer's practical purposes, the speaker shows second-personal care, and that this sort of care plays a crucial epistemic role. However, the cases Hinchman offers in support of his view are unconvincing. Either the intuitions about them are far from clear, or, as is the case with the example just given, they have a simpler explanation: 'mild' is a context-sensitive term, so it is unsurprising that it would be correct to assert that something is mild in one context and not another.

Admittedly, these brief objections do not do full justice to the views of those who have defended these strategies, and I do not intend them as decisive. Indeed, I would be happy if the objections can be answered, and there is more than one way to respond to Lackey's challenge. I present them here in this brief form to explain what motivates me to seek a different response.

3. TESTIMONIAL KNOWLEDGE AND THE AIM OF ASSERTION

In what follows, I shall argue that a second-personal interaction is required for a person to be warranted in forming a belief just on the basis of another's say-so while regarding her as a person. Hence, such an interaction is required to acquire knowledge in a distinctively testimonial way. This section presents the first part of the argument, defending the view that to yield direct warrant a source of information must have the right aim. The second part of the argument is presented in the next section, where I argue that this condition is in tension with the personal nature of testimonial exchanges, and that only a second-personal transaction can resolve this tension.

Here is the principle I will be arguing for:

Right Aim Condition (RAC): *S* is warranted in forming (or maintaining) a belief just on the basis of a deliverance *D* only if the aim of *D* is to provide true information to *S*.

Some clarifications are in order.¹⁶⁵ First, this principle states a general necessary condition on direct warrant, that is, warrant that a person obtains without needing to engage in inferential reasoning, but rather just on the basis of the deliverances of privileged epistemic

¹⁶⁵ The clarifications have been made in connection with similar principles in earlier chapters.

sources. Hence it is meant to apply not just to testimonial deliverances, but also to the deliverances from other sources traditionally thought to offer such warrant, like perception and memory. Second, to say that *S* is warranted in forming (or maintaining) a belief just on the basis of a deliverance should be taken to mean: ¹⁶⁶ if *S* forms the belief just on the basis of this source, and the right environmental conditions obtain, *S* acquires knowledge thereby. Finally, the sense in which I speak of a ‘governing aim’ may be sub-personal. It includes the sense in which scientists regularly speak of functions: the function of perception is (among others) to yield true information; the function of the heart to pump blood. To say that it is a ‘governing aim’ is to say that we evaluate these performances for correctness in terms of the extent to which they fulfill such aims: perception fails if it conveys false information; the heart fails if it does not pump blood.

In the case of testimonial exchanges, a speaker makes information available by means of an assertion. The specification of (RAC) for the case of testimony is therefore:

Right Aim for Assertion (RAA): *S* is warranted in forming a belief that *p* just on the basis of someone’s assertion that *p* only if the governing aim of the assertion is to deliver true information to *S*.

I shall presently argue for (RAA), but I should first note that the view that assertion aims at the transmission of true belief is widely held. For instance, Michael Dummett argues that “[a] man makes an assertion if he says something in such a manner as deliberately to convey the impression of saying it with the overriding intention of saying something true” (1981: 300). Similarly, many philosophers hold that, as John Turri puts it, “*the point of the*

¹⁶⁶ For simplicity, I shall henceforth focus only on belief-formation, though the points I make will apply to belief-maintenance as well.

practice of assertion is to transmit knowledge” (2016: 131).¹⁶⁷ The importance of this point for a theory of testimonial knowledge is ably drawn out by Elizabeth Fricker, who thus ends up defending a principle along the lines of (RAA).¹⁶⁸ Assertion, she argues, is governed by a knowledge norm (‘assert that p only if you know that p ’), and this norm explains why it is rational to form beliefs on the basis of another’s say-so: “In asserting that P , the asserter offers her word that P to her intended audience H . She licenses H to believe that P on her say-so. How so? What enables such a feat? It is the conventionally constituted force of her speech act that, in asserting P , the speaker S purports to speak from knowledge” (Fricker 2006a: 594). Hence, for Fricker an assertion provides warrant partly in virtue of the fact that its correctness is determined by whether what is said is known by the asserter, together with the fact that the speaker presents the assertion as a reason for the audience to form the belief.¹⁶⁹ Clearly, though, to present the assertion in this way is to present the assertion as true. On the plausible assumption that such presentation is what sets the aim of assertion, this claim is a version of (RAA).

Talk of epistemic norms of assertion might be misleading insofar as it suggests that these norms apply to assertions in general. However, not all assertions have an alethic governing aim. When one asserts something sarcastically, the assertion does not aim at

¹⁶⁷ Defenders of this view include Williamson (2000: 267), Watson (2004), and Sosa (2015, ch.8), among many others. Williamson and Turri offer this claim as an explanation of why there should be a speech act governed by a knowledge norm of assertion. None of these philosophers, to my knowledge, defends an interpersonal account of testimonial knowledge. I should note that I am partial to the view that assertion aims at transmitting knowledge rather than truth, but for dialectical purposes I state (RAA) in terms of the weaker claim.

¹⁶⁸ Of course, as a reductionist, Fricker denies that someone’s testimony can ever be the sole epistemic basis for a belief

¹⁶⁹ With Fricker, I take assertion to essentially involve an audience; in the limit case (e.g. of writing in a diary, if we wish to count these as genuine assertions) the audience might just be oneself.

delivering truth; it aims at ridicule. Similarly, when one tells a fictional story the assertion does not aim at delivering truth; perhaps it aims at entertainment. Hence, neither sarcastic assertions, nor fictional tellings are governed by anything like a knowledge norm. Consequently, by (RAA), neither sarcastic assertions nor fictional tellings could be the proper ground for testimonial beliefs.

Why hold that to be the proper ground of testimonial warrant an assertion must meet (RAA)? One reason is that it explains our intuitive judgments about certain cases.

Consider:

Mere Story: Ann is the host of a podcast that gathers stories from around the world to tell on her station. Often the stories are true, but they are sometimes merely fictional. Today Ann tells a story that she learnt from a reliable source, but for dramatic purposes she begins by saying that the story may be fictional, even though it is true, as Ann knows. Bob tunes in halfway into the show, as Ann is speaking of a zoo in South America that exhibits ligers, a cross between a lion and a tiger. Taking the story as genuine, Bob forms the belief that there is a zoo in South America that exhibits ligers.

True Story: Cesc is the host of a podcast that gathers stories from around the world to tell on his station. Often the stories are true, but they are sometimes fictional. Today Cesc tells a story that he learnt from a reliable source, so he begins by saying that this is a true story, as he knows it is. Diana tunes in halfway into the show, as Cesc is speaking of a zoo in South America that exhibits ligers, a cross between a lion and a tiger. Taking the story as genuine, Diana forms the belief that there is a zoo in South America that exhibits ligers.

The two cases are alike in all respects except that in **True Story**, Cesc's assertions have the aim of conveying knowledge, whereas in **Mere Story** Ann's assertions do not have this aim. However, it is intuitive to think that Diana's belief constitutes knowledge, whereas Bob's does not. An attractive explanation of the disparity in judgment appeals to (RAA): Diana knows and Bob does not, because her beliefs were formed on the basis of assertions that aimed at conveying knowledge whereas Ann's were not. The fact that (RAA) elegantly explains the disparity between the cases provides good grounds to hold the principle.

Here is a possible alternative explanation of the disanalogy between these cases: the reason that Diana acquires knowledge and Bob does not is that her beliefs are formed on the basis of assertions that *express* knowledge, whereas Bob's are not. However, this seems to me to either generate the wrong verdicts, or not to be a genuine alternative to the explanation in terms of governing aims. On a certain understanding of what expressing knowledge amounts to, we could imagine that Ann's assertions do indeed express knowledge: for instance, we could imagine that in telling the tale she is aiming to stick as close as possible to the testimony she heard from the person who told her the story. Even so, because she fails to represent herself as conveying knowledge, it does not seem as though her assertions could be a proper ground of belief, so they do not offer proper warrant. On the other hand, we could understand 'expressing knowledge' as 'representing oneself as conveying what is known'. But this is not a genuine alternative to the explanation proposed, because in that case Ann's assertions would have the transmission of truth as their governing aim.¹⁷⁰

It is thus hard to think of an explanation for the disanalogy between these two cases that does not appeal to (RAA), and this gives us a good reason to hold the principle. In closing, it is worth noting that (RAA) is derivable from the account of epistemic warrant defended in Chapter 4. According to that account a belief is in good epistemic standing only if it is competently formed (and maintained) relative to the aim of forming true beliefs. However, to be competently held in this way, a belief must be guided by the aim of achieving true beliefs. Now, when we form beliefs on the basis of the perceptual system we rely on this system to provide us with true information for the formation of belief, and

¹⁷⁰ Arguably, they would have a stronger aim: the transmission of knowledge.

when we form a belief on the basis of someone else's say-so, we rely on their communicative act as a source of true information for the formation of belief. As such, these are joint performances, the first between ourselves as doxastic agents and a sub-personal system, the second between ourselves as doxastic agents and another person as an informant. If these joint performances are to provide us with epistemic warrant, therefore, they must aim at providing true information on the basis of which we can form true beliefs. If they didn't have this aim, then, even when a true belief results from such deliverances, as in **Mere Story**, the fact that the belief is true is accidental relative to the belief-forming process. Hence, the account of warrant from Chapter 4 entails that to provide us with direct warrant a source must aim at providing true information, as per (RAC).

4. THE EPISTEMIC FUNCTION OF ASSURANCE

There are good reasons, then, to hold that to be the proper grounds of testimonial belief, assertions need to aim at transmitting truth. Because assertions often meet this requirement, it would seem as though, like our perceptual systems, other persons can through their assertions give us proper warrant for belief just on the basis of those assertions. As we are about to see, however, focusing on the case of assertions generates a difficulty that does not arise for the case of perception.

Here is the problem. In the case of exercises of the perceptual system, the right aim condition obtains naturally: it is the function of our perceptual systems to deliver correct information so that we might form true beliefs about the world. However, other persons do not have that function in relation to our beliefs, and this raises an epistemic difficulty. As Angus Ross (1986) writes:

The main problem with the idea that the hearer views the speaker's words as evidence arises from the fact that, unlike the examples of natural signs which spring most readily to mind, *saying something is a deliberate act under the speaker's conscious control and the hearer is aware that this is the case . . .* If a speaker's words are evidence of anything, they have that status only because he has chosen to use them [my emphasis]. (p.77)

Building on this theme, Moran (2005) notes:

When I learn of someone's beliefs through what they tell me, I am dependent on such things as their discretion, sincerity, good intentions—in short, on how they deliberately present themselves to me—in a way that I am not dependent when I infer their beliefs in other ways. People are known to lie, exaggerate, and otherwise speak in ways that do not express their genuine beliefs. (p.5)

The difficulty that Ross and Moran are alerting us to stems from the fact that assertions are intentional actions, and the speaker can thus do with them what they will. As such, assertions cannot simply be assumed to have an alethic aim, as would be required by (RAA) if they are to be the proper source of beliefs. Indeed, in light of the above considerations, Moran's focus on cases of deception is actually somewhat misleading. For it is not just that speakers can use speech to mislead us, in such a way that forming beliefs on the basis of what they say would lead to largely false beliefs; it is rather that speakers can use assertoric speech for any number of purposes, not just to deceive, but also for much more harmless purposes like telling fictional stories, make jokes, persuade, thrill, raise admiration, or console. Unlike exercises of our perceptual systems that necessarily have an epistemic aim, informing others is just one among many things that we can aim to do with assertoric speech.

To recognize this—to recognize that the speaker is capable of having myriad aims in what she does, including what she does with her words—is constitutive of recognizing the other as a person, a being free to set her own aims for her actions. However, it was argued that one could be warranted in forming beliefs just on the basis of another's say-so just in case the person's assertions are governed by the aim of transmitting true information.

A listener thinking about whether to form a belief just on the basis of what another says would thus seem to face a dilemma: she can either regard the speaker as a person, in which case she will not be entitled to form beliefs just on the basis of what she says (since this will require her to regard the speaker as not having a determinate aim in her assertions); alternatively, she can regard the speaker as a mere instrument to the truth, a system whose function is to transmit true information, in the way her perceptual system does.

Note that this dilemma does not call into question the possibility of acquiring knowledge on the basis of another's say-so. The second horn is compatible with this option: a listener can acquire such knowledge if she takes the speaker as having a determinate aim (and the speaker does have that aim). What the dilemma calls into question is rather the possibility of acquiring knowledge from a person while regarding her as such. I am taking this possibility as constitutive of the notion of testimonial knowledge. Thus, I take the dilemma to call into question the possibility of testimonial knowledge.¹⁷¹

Moreover, we should distinguish the foregoing problem from a different, skeptical problem with which it might easily be confused. The skeptical problem is about how we can ever be *certain* that another person is being sincere, reliable, or telling the truth. This, inasmuch as it is a real problem, is not one that affects testimony in particular. Similar concerns could be raised about perception and memory, for instance. The problem that I'm interested in arises only for testimony, and it arises because of the particular nature of persons as beings capable of setting their own aims. Because persons can set their own aims, it seems impossible to relate to someone as a person while at the same time relating

¹⁷¹ I am presupposing that to acquire knowledge from a certain source, the subject must be able (if she has the conceptual means) to regard herself as being rational in relying on that source. (This is a much weaker requirement than the requirement that she must rely herself as being rational in so relying).

to them as a source of direct warrant, because to do the latter requires us to regard the other as someone whose aim is the very determinate one of providing us with true information.

With this in mind, consider a simple potential solution to the problem: the speaker and the hearer should come to an agreement to the effect that the hearer can rely on and form beliefs on the basis of what the speaker says, because the speaker agrees to only transmit true information. However, this only pushes the problem back. For how could two people ever come to such an agreement without already being able to trust each other in the characteristic way of testimonial knowledge? Wouldn't the speaker have to say something to the effect of, 'You can rely on me when it comes to such and such' or 'This is what I'm going to do'? If so, the problem simply re-emerges, for we can then ask: what reason has the person to trust these assertions, given that they can be governed by non-epistemic aims?¹⁷²

To find a more satisfying solution to the problem it is worth noting that a parallel problem to this one emerges in the practical realm. Suppose Ann wants to rely on Bob to Φ (given some practical aim that she has). How can she do that? It seems she cannot rely on Bob to Φ while regarding him as a person given that, as a person, Bob is free to do as he pleases. Hence, it seems the only way to obtain the relevant reliance is to stop treating Bob as a person. For instance, Ann could force Bob to Φ (through threat or force, for instance), or take advantage of the fact that Bob regularly Φ s in the way in which one might take advantage for one's gardening of the fact that the sun regularly comes up.¹⁷³ Yet,

¹⁷² I am here essentially echoing Baier (1986), who presents a sustained criticism of the view that the notion of a contract is more basic than the notion of trust. Among the many reasons that Baier cites in support of the opposite direction of dependence is that children cannot form contracts, but they can certainly trust.

¹⁷³ As Steve Darwall pointed out, Ann could treat Bob as a person by offering him incentives to do as she wants him to do, incentives that she believes Bob will (rationally) accept. However, this suffers from the same problem that the agreement strategy encounters: the view appeals to an offer by Ann, but an offer of

neither of these would be a form of personal reliance which seems now a paradoxical notion: the nature of persons as free beings, and the nature of reliance as requiring others to act in determinate ways, seem to clash.¹⁷⁴

The parallel with the practical case, however, suggests a third possibility: Ann can rely on Bob, while regarding him as a person, if Bob promises to her that he will Φ .¹⁷⁵ For by promising that he will Φ , Bob thereby freely binds himself to Φ , giving Ann a distinctive reason to rely on him to Φ (if she accepts the promise), a second-personal reason that is fundamentally grounded in the very act of promising. Because promises are freely given, but also, by their very nature, set aims for the promisor to which she is bound, they enable the promisee to rationally rely on the promisor while regarding her as a person. Importantly, these standards of correctness hold just in virtue of the act of promising, and the authority over those standards is transferred to the promisee. Thus, whereas an agent who decides to Φ can correctly not- Φ by simply giving up that aim, a person who promises to Φ is bound by that aim even if she changes her mind. The only way in which she can now not- Φ correctly is if she is released from the promise by the promisee. This is the transfer of authority that is characteristic of second-personal exchanges.

It is also crucial that the commitment to an aim that promises give rise to is internal to the act of promising: to promise to Φ just is, among other things, to commit oneself to

this sort is just a sort of promise, and what we are trying to make sense of is how it could be rational to accept a promise in the first place.

¹⁷⁴ The two forms of reliance correspond to Darwall's (2017) distinction between an *expectation-of* and an *expectation-that*. As in the case of trust, the possibility of expecting-that is not called into question here; only the possibility of expecting-of.

¹⁷⁵ Many authors have explored the parallels between testimonial and promissory exchanges. See e.g. Brandom (1983), Watson (2004), Fricker (2006a: 602-03), Friedrich and Southwood (2011), and Marušić (2015).

Φ ing and to bind oneself to that standard. If such a commitment were external to the act, so that one could promise to Φ without binding oneself to the aim of Φ ing, the same problem that we encountered with the agreement strategy as a ground of trust would reappear in this case. In order to rationally rely on another's promise to Φ , the promisee would need a further assurance from the promisor, since the promise could be given without providing such assurance. Clearly, the needed assurance could not be attained through a second-order promise to be bound by the first-order promise to Φ , at least on this understanding of promises. For the same worries would arise for this second-order promise, giving rise to a vicious regress. Hence, the assurance that promises give must be internal to them.

Promises solve a practical problem: the problem of how one person can rationally rely on another as a person, that is, while regarding the other as free to do as she wills. They solve this problem in virtue of their nature as acts that set their own norms of correctness. Hence, when an agent makes a promise, she must freely bind herself to such a standard. What I want to suggest is that tellings have a similar structure, which enables them to solve the parallel epistemic problem above. When a person tells someone something, she thereby freely sets as the standard of her (linguistic) behaviour the deliverance of true information to the listener. To tell someone something is just to undertake an act with this normative structure, whereby a speaker freely forsakes authority to set her own aims for what she says. As such, when a speaker says something to a listener, the speaker has a reason to communicate the truth, and the listener has a reason to believe what the speaker says, and both reasons are grounded in the act of telling itself.¹⁷⁶ Because

¹⁷⁶ Of course, there might be cases where the listener misidentifies the act in question, as in **Mere Story**. The same can happen with promises, or, indeed, any action whatsoever, which can be misinterpreted. This does

tellings are freely undertaken acts, but also set standards of correctness for the speaker's assertion, they enable the hearer to rationally trust the speaker and form a belief just on the basis of what she says while treating her as a person.

This parallel structure of promises and tellings is precisely the one that was identified in Section 2 as characteristic of second-personal reasons. When a promisee accepts a promise, she has a reason to rely on the promisor to act in a certain way that is fundamentally grounded in the latter's promise. Asked why she is relying on the other, she should say, 'because she promised'. Likewise, when a listener accepts testimony, by trusting a speaker, she has a reason to believe what the speaker says that is fundamentally grounded in the latter's telling. Asked why she believes what the other says, she should say, 'because she told me so'.¹⁷⁷

In the previous section I explained why the performances on the basis of which we form beliefs must have the right aim to provide us with direct warrant for belief. For reasons familiar from Chapter 4, it is clear at any rate that mere reliability is not enough to obtain such warrant. In the case of testimonial transactions, understood as joint performances, the right aim must govern the actions of both the speaker and the listener. Since a second-personal transaction is required for these performances to have the right aim in the case of testimony, it is clear why such a transaction is needed. The account thus explains how

not pose a problem to the position I am advocating, which aims not at securing an infallible way to attain testimonial knowledge, but rather just to illuminate the conditions that make it possible for us to attain such knowledge in the cases in which we do attain it.

¹⁷⁷ I take it that by believing her, the listener is thereby relying on her, that is, for her epistemic performance. I therefore disagree with Marušić (2017), who holds that trust is more primitive than reliance. Still, I agree with him that reliance is reliance on one's performance; but in light of the virtue-theoretic treatment of belief in the previous section, I would take it that forming (and maintaining) a belief is among the performances that can be done on the basis of reliance. The fact that certain kinds of reasons are not proper for belief-formation is explained by the fact that this is an epistemic performance, with an alethic constitutive aim.

testimonial exchanges can give us direct warrant to form beliefs, without giving up on the distinctively personal nature of testimonial knowledge. It also offers an explanation of how there could be an epistemic difference between forming a belief on the basis of being told something and merely overhearing it. For the account requires that the speaker and the listener be engaged in a joint performance, and mere overhearers do not engage in a joint performance with the speaker.¹⁷⁸

Noticing the crucial role played in the story by the second-personal structure of tellings also provides a response to Lackey's challenge. To recall, the challenge takes the form of a dilemma with respect to the conditions that a testimonial belief must meet to constitute knowledge: either non-second-personal factors—like truth and reliability—need to be met, or they do not. If they do, the non-personal factors seem to do all the epistemic work; if they do not, it is not clear why the belief would be in good epistemic standing such as to constitute knowledge. Either way, second-personal factors seem to play no epistemic role.

As I noted above, the argument has an unstated assumption, namely, that the epistemic conditions that we would otherwise be inclined to think are required for proper epistemic standing can obtain independently of whether the second-personal factors obtain.

¹⁷⁸ In some cases, the context makes it clear that anyone in the audience will be suitably related to the speaker, e.g. when someone is giving a speech, though the information is not directed at anyone in particular, any listener in the audience will be engaged in a joint activity with the speaker. They thus count as 'being told' something by my account, even if not by natural language standards of 'being told'. By contrast, it is an implication of my account that someone who reads a diary without permission cannot gain testimonial knowledge, since the information was not intended to be transmitted to them. Some have taken such implications to favour a non-interpersonal account (e.g. Weiner (ms)), but I do not see why the diary case should be treated in the same way as paradigmatic cases of testimonial knowledge acquisition (indeed, I incline to the opposite view, though my intuitions here are arguably infected by theory). And as I continue to emphasize, that one cannot acquire *testimonial* knowledge from a diary is compatible with holding that one can acquire (non-testimonial) knowledge on that basis, which is obvious one can.

However, if the view advocated here is correct, this assumption is false: we can obtain the specific kind of warrant that a testimonial source gives us—a warrant that is both direct and personal—only on the basis of a second-personal transaction. Such a transaction is needed for one to be warranted in forming a belief just on the basis of what another person says, while regarding her as a person. Since the ability to regard the other as a person is constitutive of testimonial transactions, only on the basis of a second-personal exchange could a person get the specific kind of warrant that testimony makes available. As such, second-personal features play a crucial epistemic role in the acquisition of testimonial knowledge.

5. EVIDENTIALISM AND EPISTEMIC PARTIALITY

The argument I have presented for a second-personal account of testimonial knowledge appeals to a tension between two general features, distinctive of this kind of knowledge. I have argued that this tension between them can only be resolved through a second-personal transaction. Unlike other defenders of interpersonal accounts, therefore, I have placed little emphasis on cases of epistemic partiality, like **Accused Friend** and **Quitting Addict**. One reason for this has already been given: I do not think these cases are dialectically strong to appeal to, because the opponent of interpersonal views can reject them as ones where non-epistemic reasons outweigh epistemic reasons. Nevertheless, I think these cases are important and interesting, so I shall close by examining what light the account I have presented might shed on them. This discussion will also clarify the sense in which the account I have presented is non-evidential.

First, with regards to cases of epistemic partiality such as **Accused Friend** and **Quitting Addict**, I agree with other defenders of interpersonal views of testimonial knowledge that in these cases we are epistemically permitted to believe against the evidence. That is, if ‘belief against the evidence’ is understood to mean that we can believe what the person tells us (they’re innocent, or they will quit), even if independently available evidence overwhelmingly suggested that what they say is false. Given such a set-up, we are permitted to believe what the speaker tells us, because if the other person competently checks and delivers the information, that will suffice for our belief to be warranted. On the other hand, there is a broader understanding of evidence where it just means ‘epistemic reason’, and in that sense I would say that testimony gives us evidence of a particular sort, one meant to bracket other forms of evidence, such as statistical evidence.

To explain what this means, recall the example from Rowe (1979) discussed in Chapter 5. A ship has suffered a terrible accident of a sort that makes it overwhelmingly likely that no one in the ship survived. On the basis of this evidence, it would be wrong to judge that anyone survived. However, suppose you were on the ship, and you miraculously survived. You would then be aware of your survival, in which case it would be wrong for you to think that you didn’t survive or to give weight to considerations that would lead to that conclusion: your awareness of surviving should bracket other considerations that would tell against your surviving. My suggestion is that since testimony is a basic knowledge source, it gives us reasons of this bracketing kind. These reasons entitle us to ignore other considerations that might tell against our beliefs formed on their basis.

One might question this view on the grounds that rationality demands that we always weigh *all* the considerations that are readily available pertaining to the question of

what to believe. On this view, setting aside any considerations would therefore amount to irrationality. However, it is far from clear that rationality demands so much from us, and there is reason to think the opposite, that sometimes rationality demands setting some considerations aside.

One area where this is clear is in the realm of practical authority. As Joseph Raz has persuasively argued (1985a, 1981, 1985b, 2010), one of the marks of authority is that the reasons it gives rise to are exclusionary: authority gives us reasons not only to proceed in a certain way, but to set aside other considerations relevant to the matter at hand. Thus, for instance, if you were a sailor in a ship, and your captain commanded you to clean the main cabin, you should do so, even if you think that your time would be spent more efficiently doing something else. As Raz argues, part of the reason why it is rational to do this is that one of the primary functions of authority is precisely to ‘pass the buck’ of decision making on to another person. To be an authority is to occupy a position where others can rely on you to weigh the relevant considerations and come to a decision that is final. It would thus undermine the very purpose of having an authority if those who are supposed to follow the commands took matters in their own hands and attempted to resolve the matter on their own.

In cases of authority we rely on others to weigh practical reasons so as to make certain decisions. In cases of testimony we rely on others to collect true information so as to make it possible for us to form true beliefs and acquire knowledge on that basis. And just as the point of relying on authorities would be undermined if we went about making our own decisions when an order has been given, the point of testimonial reliance would be undermined if we went about collecting our own information when someone has told us

something.¹⁷⁹ Of course, in both cases it is possible to question whether we should be reliant: one can have grounds to question both authority and trustworthiness, and this raises further complications.

One complication arises from the fact that in cases of epistemic partiality, as standardly presented, the speaker is a friend, or someone close to the listener. In light of that, there is some pressure to think that we are not only permitted to trust them, but required to do so, despite the existence of strong evidence against what they say. After all, doing otherwise would amount to questioning the trustworthiness of the speaker. Still, this obligation could well arise from norms of friendship, rather than epistemic norms, which might lead us to think of these cases as ones where these two norms clash. On the other hand, if the account of testimonial warrant I defended is correct, there may not be a conflict even in these cases, at least in the ‘good case’, where the friend is competently and sincerely telling the truth. For if the speaker is sufficiently competent at checking and delivering the truth, then the belief formed on its basis will be warranted, and well-formed by epistemic standards as well.

Yet, this verdict depends on taking a stand on a second issue that raises a further difficulty: in cases like **Accused Friend** the independent evidence might be taken as a good reason to question the friend’s trustworthiness. In other words, these cases might seem to be ones where you have good reason to question the source on the basis of which you would form the belief; but then the belief thus formed would inevitably be defeated. Take, then, a case where the speaker is competently and sincerely telling the truth, but where

¹⁷⁹ My view here is influenced by Zagzebski’s (2015) pioneering work on the relationship between authority and epistemology.

there is also strong independent evidence to think that she must be saying something false. I am inclined to think that even under these circumstances, a belief could be warranted and constitute knowledge, even if you are not in a position to discriminate between this case and one where she is lying or has been misled, where you would not be so warranted;¹⁸⁰ for it is the obtaining of the relevant external conditions that justify you in setting aside conflicting considerations. However, I am far from confident in my judgment about this case.

The point I want to insist on is that, inasmuch as independent considerations might tell against believing what the friend says, they would have to do so indirectly by way of questioning the trustworthiness of the speaker, rather than directly, by presenting opposing evidence against the claim under consideration. For testimonial reasons, like authoritative reasons, are exclusionary reasons. Hence, one cannot take something as a testimonial reason and weigh it against contrary evidence. Such evidence might lead you to not take it as a testimonial reason at all, but if you do take the reason as testimonial, that reason will settle the matter for you, and you will be rationally required to set aside evidence that conflicts with it as misleading.

¹⁸⁰ My views here are influenced by Lasonen-Aarnio (2010).

References

- Adler, Jonathan. 2008. 'Epistemological problems of testimony.' in, *Stanford Encyclopedia of Philosophy*.
- Anscombe, G. E. M. 1958. *Intention* (Basil Blackwell: Oxford).
- . 1963. 'The two kinds of error in action', *Journal of Philosophy*, 60: 393-401.
- . 1981 [1965]. 'Thought and Action in Aristotle: What is 'Practical Truth'?' in, *From Parmenides to Wittgenstein* (University of Minnesota Press: Minneapolis).
- Anscombe, Gertrude Elizabeth Margaret. 1979. 'What is it to Believe Someone?' in C. F. Delaney (ed.), *Rationality and Religious Belief* (University of Notre Dame Press).
- Austin, John L. 1975. 'How to do things with words, ed. JO Urmson and Marina Sbisá', *Cambridge: Harvard University Press*, 120: 125.
- Baier, Annette. 1986. 'Trust and antitrust', *Ethics*, 96: 231-60.
- Baier, Annette C. 1970. 'Act and intent', *Journal of Philosophy*, 67: 648-58.
- Baker, Judith. 1987. 'Trust and Rationality', *Pacific Philosophical Quarterly*, 68: 1.
- Baker, Judith, and Philip Clark. 2017. 'Epistemic buck-passing and the interpersonal view of testimony', *Canadian Journal of Philosophy*: 1-22.
- Bennett, Jonathan. 1988. *Events and their Names* (Oxford University Press: Oxford).
- Berker, Selim. 2008. 'Luminosity Regained', *Philosophers' Imprint*, 8: 1-22.
- Blackson, Thomas A. 2007. 'On Williamson's Argument for (I) in His Anti-Luminosity Argument', *Philosophy and Phenomenological Research*, 74: 397-405.
- Boghossian, Paul. 2014. 'What is inference?', *Philosophical Studies*, 169: 1-18.
- BonJour, Laurence. 1985. *The Structure of Empirical Knowledge* (Harvard University Press).
- Brandom, Robert. 1983. 'Asserting', *Noûs*, 17: 637-50.
- Bratman, Michael. 1999 [1987]. *Intention, Plans, and Practical Reason* (The David Hume Series: CSLI Publications).
- Brueckner, Anthony, and M Oreste Fiocco. 2002. 'Williamson's anti-luminosity argument', *Philosophical Studies*, 110: 285-93.
- Burge, Tyler. 1993. 'Content preservation', *The philosophical review*, 102: 457-88.
- . 1997. 'Interlocution, perception, and memory', *Philosophical Studies*, 86: 21-47.
- . 2003. 'Perceptual entitlement', *Philosophy and Phenomenological Research*, 67: 503-48.
- . 2010. *Origins of objectivity* (Oxford University Press).
- Burge, Tyler, and Christopher Peacocke. 1996. "Our Entitlement to Self-Knowledge: I. Tyler Burge." In *Proceedings of the Aristotelian Society*, 91-116. JSTOR.
- Charles, David. 2018. 'Processes, Activities and Actions.' in Rowland Stout (ed.), *Process, action, and experience* (Oxford University Press).
- Chisholm, Roderick M. 1987. 'The truths of reason'.
- Coady, C. A. J. 1992. *Testimony: A Philosophical Study* (Oxford University Press).
- Cohen, Stewart. 1984. 'Justification and truth', *Philosophical Studies*, 46: 279--95.
- . 2010. 'Luminosity, Reliability, and the Sorites', *Philosophy and Phenomenological Research*, 81: 718-30.
- Colaço, David, Wesley Buckwalter, Stephen Stich, and Edouard Machery. 2014. 'Epistemic intuitions in fake-barn thought experiments', *Episteme*, 11: 199-212.

- Cresto, Eleonora. 2012. 'A defense of temperate epistemic transparency', *Journal of philosophical logic*, 41: 923-55.
- Darwall, Stephen. 2001. 'Because I Want It', *Social Philosophy and Policy*, 18: 129-53.
- . 2006. *The second-person standpoint: Morality, respect, and accountability* (Harvard University Press).
- . 2017. 'Trust as Second-Personal Attitude (of the Heart).' in Paul Faulkner and Thomas Simpson (ed.), *The Philosophy of Trust* (Oxford University Press: Oxford).
- Davidson, Donald. 1973. 'Freedom to act.' in Ted Honderich (ed.), *Essays on Freedom of Action* (Routledge).
- . 2001a. 'Agency', *Essays on actions and events*: 43-61.
- . 2001b. 'Intending', *Essays on actions and events*: 83-102.
- de Saint-Exupéry, Antoine. 2015. *Le petit prince: avec des aquarelles de l'auteur* (Ernst Klett Sprachen GmbH).
- DeRose, Keith. 1995. 'Solving the skeptical problem', *The philosophical review*, 104: 1-52.
- . 2002. 'Assertion, knowledge, and context', *The philosophical review*, 111: 167-203.
- Descartes, Rene. 2013 [1641]. *Meditations on First Philosophy: With selections from the objections and replies*.
- Donnellan, Keith S. 1963. 'Knowing what I am doing', *Journal of Philosophy*, 60: 401-09.
- Douven, Igor. 2006. 'Assertion, knowledge, and rational credibility', *The philosophical review*, 115: 449-85.
- Dummett, Michael. 1981. *Frege: Philosophy of language* (Harvard University Press).
- Egré, Paul. 2006. 'Reliability, margin for error and self-knowledge.' in D.H. Pritchard and V. Hendricks (eds.), *New waves in epistemology* (Palgrave-Macmillan).
- Egré, Paul, and Jérôme Dokic. 2008. 'Margin for Error and the Transparency of Knowledge', *Synthese*: 1-36.
- El-Alayli, Amani, Ashley A Hansen-Brown, and Michelle Ceynar. 2018. 'Dancing Backwards in High Heels: Female Professors Experience More Work Demands and Special Favor Requests, Particularly from Academically Entitled Students', *Sex Roles*: 1-15.
- Falvey, Kevin. 2000. 'Knowledge in intention', *Philosophical Studies*, 99: 21-44.
- Faulkner, Paul. 2007. 'On Telling and Trusting', *Mind*, 116: 875-902.
- . 2011. *Knowledge on trust* (Oxford University Press Oxford).
- Ferrero, Luca. 2009. 'Conditional Intentions', *Noûs*, 43: 700 - 41.
- . 2017. 'Intending, acting, and doing', *Philosophical Explorations*, 20: 13-39.
- Fleming, Brice Noel. 1964. 'On intention', *The philosophical review*, 73: 301-20.
- Foot, Philippa. 1967. 'The problem of abortion and the doctrine of double effect', *Oxford Review*, 5: 5-15.
- Frankfurt, Harry G. 1978. 'The Problem of Action', *American Philosophical Quarterly*, 15: 157-62.
- Fricke, Elizabeth. 1994. 'Against Gullibility.' in A. Chakrabarti and B. K. Matilal (eds.), *Knowing from Words* (Kluwer Academic Publishers).
- . 1995. 'Critical notice: Telling and trusting: Reductionism and anti-reductionism in the epistemology of testimony', *Mind*, 104: 393-411.

- . 2004. 'Testimony: Knowing through being told.' in M. Sintonen, J. Wolenski and I. Niiniluoto (eds.), *Handbook of Epistemology* (Kluwer Academic Publishers).
- . 2006a. 'Second-hand knowledge', *Philosophy and Phenomenological Research*, 73: 592–618.
- . 2006b. 'Varieties of Anti-Reductionism About Testimony? A Reply to Goldberg and Henderson', *Philosophy and Phenomenological Research*, 72: 618-28.
- Fricker, Elizabeth, and David E. Cooper. 1987. 'The Epistemology of Testimony', *Aristotelian Society Supplementary Volume*, 61: 57 - 106.
- Fricker, Miranda. 2007. *Epistemic injustice: Power and the ethics of knowing* (Oxford University Press).
- Friedrich, Daniel, and Nicholas Southwood. 2011. 'Promises and Trust.' in Hanoch Sheinman (ed.), *Promises and Agreement: Philosophical Essays* (Oxford University Press).
- Gibbons, John. 2010. 'Seeing What You 're Doing'.' in T. Szabo Gendler and J. Hawthorne (eds.), *Oxford Studies in Epistemology* (Oxford University Press).
- Goldberg, Sanford. 2006. 'Reductionism and the distinctiveness of testimonial knowledge.' in Jennifer Lackey and Ernest Sosa (eds.), *The Epistemology of Testimony* (Oxford University Press).
- Goldman, Alvin. 1976. 'Discrimination and perceptual knowledge', *The Journal of Philosophy*, 73: 771-91.
- Goldman, Alvin I. 1970. *A Theory of Human Action* (Princeton University Press).
- Graham, Peter J. 2012. 'Epistemic entitlement', *Noûs*, 46: 449-82.
- . 2014. 'Functions, Warrant, History.' in Abrol Fairweather and Owen Flanagan (eds.), *Naturalizing Epistemic Virtue* (Cambridge University Press).
- Greco, Daniel. 2014. 'Could kk be ok?', *The Journal of Philosophy*, 111: 169-97.
- . 2015. 'Iteration Principles in Epistemology II: Arguments Against', *Philosophy Compass*, 10: 765-71.
- . forthcoming. 'Justifications and Excuses in Epistemology', *Noûs*.
- Greco, John. 2003. "Knowledge as Credit for True Belief". in Michael DePaul and Linda Zagzebski (eds.), *Intellectual Virtue: Perspectives From Ethics and Epistemology* (Clarendon Press).
- Grice, H. P. 1971. 'Intention and Uncertainty', *Proceedings of the British Academy*, 57: 263-79.
- . 1988. 'The causal theory of perception.' in Jonathan Dancy (ed.), *Aristotelian Society Supplementary Volume* (Oxford University Press).
- Hampshire, Stuart. 1983. *Thought and Action* (University of Notre Dame Press).
- Hanrahan, Rebecca Roman, and Louise M. Antony. 2005. 'Because I said so: Toward a feminist theory of authority', *Hypatia*, 20: 59-79.
- Harman, Gilbert. 1986a. *Change in View* (MIT Press).
- . 1986b. 'Willing and Intending.' in R. E. Grandy and R. Warner (ed.), *Philosophical Grounds of Rationality* (Oxford University Press: Oxford).
- . 1997. 'Practical reasoning', *Review of Metaphysics*, 29: 431--63.
- Hawthorne, John. 2004. *Knowledge and lotteries* (Oxford University Press).
- Henry, Devin, and Karen Margrethe Nielsen. 2015. *Bridging the Gap Between Aristotle's Science and Ethics* (Cambridge University Press).

- Hieronymi, Pamela. 2006. 'Controlling attitudes', *Pacific Philosophical Quarterly*, 87: 45-74.
- Hill, Christopher S., and Joshua Schechter. 2007. 'Hawthorne's lottery puzzle and the nature of belief', *Philosophical Issues*, 17: 120-22.
- Hinchman, Edward. 2005. 'Telling as inviting to trust', *Philosophy and Phenomenological Research*, 70: 562–87.
- . 2014. 'Assurance and warrant', *Philosophers' Imprint*, 14: 1-58.
- Holton, Richard. 1994. 'Deciding to trust, coming to believe', *Australasian Journal of Philosophy*, 72: 63 – 76.
- . 1999. 'Intention and Weakness of Will', *Journal of Philosophy*, 96: 241.
- . 2009. *Willing, Wanting, Waiting* (Oxford University Press UK).
- Hornsby, Jennifer. 1980. *Actions* (Routledge and Kegan Paul).
- . 2016. 'Intending, knowing how, infinitives', *Canadian Journal of Philosophy*, 46: 1-17.
- Horvath, Joachim, and Alex Wiegmann. 2015. 'Intuitive expertise and intuitions about knowledge', *Philosophical Studies*: 1-26.
- Hume, David. 1904. *Enquiry Concerning Human Understanding* (Clarendon Press).
- . 2007. "A Treatise of Human Nature: A critical edition, edited by David Fate Norton, and Mary J." In.: Norton. Oxford: Clarendon Press.
- Johnston, Mark. 1992. 'How to Speak of the Colors', *Philosophical Studies*, 68: 221-63.
- Katz, Bernard D. 1978. 'Kim on events', *Philosophical Review*, 87: 427-41.
- Keller, Simon. 2004. 'Friendship and Belief', *Philosophical Papers*, 33: 329-51.
- Kim, Jaegwon. 1966. 'On the psycho-physical identity theory', *American Philosophical Quarterly*, 3: 227-35.
- . 1973. 'Causation, nomic subsumption, and the concept of event', *Journal of Philosophy*, 70: 217-36.
- . 1976. 'Events as Property Exemplifications.' in M. Brand and D. Walton (eds.), *Action Theory* (D. Reidel).
- Klein, Peter D. 2011. 'Infinitism.' in Sven Bernecker and Duncan Pritchard (eds.), *Routledge Companion to Epistemology* (Routledge).
- Lackey, Jennifer. 1999. 'Testimonial knowledge and transmission', *Philosophical Quarterly*, 50: 471-90.
- . 2007. 'Norms of assertion', *Noûs*, 41: 594-626.
- . 2008. *Learning from words: Testimony as a source of knowledge* (Oxford University Press on Demand).
- Langton, Rae. 2004. 'Intention as faith', *Royal Institute of Philosophy Supplement*, 55: 243-58.
- . 2009. 'Speaker's Freedom and Maker's Knowledge.' in, *Sexual solipsism: Philosophical essays on pornography and objectification* (Oxford University Press).
- . 2015. 'How to Get a Norm from a Speech Act', *The Amherst Lecture in Philosophy*, 10: 1-33.
- Lasonen-Aarnio, Maria. 2010. 'Unreasonable knowledge', *Philosophical Perspectives*, 24: 1-21.
- Lehrer, Keith, and Stewart Cohen. 1983. 'Justification, truth, and coherence', *Synthese*, 55: 191-207.

- Leitgeb, Hannes. 2002. 'Timothy Williamson, Knowledge and its Limits, Oxford: Oxford University Press, 2000, xi+ 340pp', *Grazer Philosophische Studien*, 65: 195-205.
- Levy, Yair. 2018. 'Why cognitivism?', *Canadian Journal of Philosophy*, 48: 223-44.
- Lewis, David. 1979. 'Attitudes de dicto and de se', *The philosophical review*, 88: 513-43.
- Littlejohn, Clayton. 2015. 'A Plea for Epistemic Excuses.' in Fabian Dorsch Julien Dutant (ed.), *The New Evil Demon Problem* (Oxford University Press).
- Marcel, Anthony J. 2003. 'The sense of agency: Awareness and ownership of action.' in Johannes Roessler and Naomi Eilan (eds.), *Agency and Self-Awareness: Issues in Philosophy and Psychology* (Oxford: Clarendon Press).
- Marušić, Berislav. 2012. 'Belief and difficult action', *Philosopher's Imprint*, 12.
- . 2015. *Evidence and agency: Norms of belief for promising and resolving* (Oxford University Press, USA).
- . 2017. 'Trust, Reliance and the Participant Stance', *Philosopher's Imprint*, 17.
- Marušić, Berislav, and John Schwenkler. 2018. 'Intending Is Believing: A Defense of Strong Cognitivism', *Analytic Philosophy*, 59: 309-40.
- McDowell, John. 1994. 'Knowledge by hearsay.' in A. Chakrabarti and B. K. Matilal (eds.), *Knowing From Words* (Kluwer Academic Publishers).
- . 2010. 'What is the Content of an Intention in Action?', *Ratio*, 23: 415-32.
- McKinnon, Rachel. 2013. 'The Supportive Reasons Norm of Assertion', *American Philosophical Quarterly*, 50: 121-35.
- . 2015. *Norms of Assertion: Truth, Lies, and Warrant* (Palgrave-Macmillan).
- McMyler, Benjamin. 2011. *Testimony, trust, and authority* (OUP USA).
- . 2013. 'The epistemic significance of address', *Synthese*, 190: 1059-78.
- Mele, Alfred R., and Paul K. Moser. 1994. 'Intentional action', *Noûs*, 28: 39-68.
- Miracchi, Lisa. 2015. 'Competence to know', *Philosophical Studies*, 172: 29-56.
- Moore, Michael S. 2010. *Act and crime: The philosophy of action and its implications for criminal law* (Oxford University Press).
- Moran, Richard. 2001. *Authority and estrangement: An essay on self-knowledge* (Princeton University Press).
- . 2004. 'Anscombe on 'Practical Knowledge'.' in J. Hyman and H. Steward (eds.), *Royal Institute of Philosophy Supplement* (Cambridge University Press).
- . 2005. 'Getting told and being believed', *Philosophers' Imprint*, 5: 1-29.
- . 2013. 'Testimony, Illocution and the Second Person', *Aristotelian Society Supplementary Volume*, 87: 115-35.
- Moss, Jessica. 2014. 'Right Reason in Plato and Aristotle: On the Meaning of Logos', *Phronesis*, 59: 181-230.
- Myers-Schulz, Blake, and Eric Schwitzgebel. 2013. 'Knowing that P without believing that P', *Noûs*, 47: 371-84.
- Neta, Ram, and Guy Rohrbaugh. 2004. 'Luminosity and the Safety of Knowledge', *Pacific Philosophical Quarterly*, 85: 396-406.
- Newstead, Anne. 2006. 'Knowledge by Intention? On the Possibility of Agent's Knowledge.' in Stephen Hetherington (ed.), *Aspects of Knowing* (Elsevier Science).
- Nozick, Robert. 1981. *Philosophical explanations* (Harvard University Press).
- O'Brien, Lucy. 2007. *Self-knowing agents* (Oxford University Press).
- O'Shaughnessy, Brian. 1980. *The will* (Cambridge University Press Cambridge).

- Pacherie, Elisabeth. 2007. 'The Sense of Control and the Sense of Agency', *PSYCHE: An Interdisciplinary Journal of Research On Consciousness*, 13: 1 - 30.
- Paul, Sarah K. 2009a. 'How We Know What We're Doing', *Philosophers' Imprint*, 9: 1-24.
- . 2009b. 'Intention, belief, and wishful thinking: Setiya on “practical knowledge”', *Ethics*, 119: 546-57.
- . 2012. 'How we know what we intend', *Philosophical Studies*, 161: 327-46.
- Paul, Sarah K., and Jennifer M. Morton. 2018. 'Believing in Others', *Philosophical Topics*, 46: 75-96.
- Peacocke, Christopher. 2003. *The Realm of Reason* (Oxford University Press).
- Pears, David. 1985. 'Intention and belief.' in Bruce Vermazen and Merrill B. Hintikka (eds.), *Essays on Davidson: Actions and Events* (Oxford University Press: Oxford).
- Piñeros Glasscock, Juan S. forthcoming. 'Reason in Action in Aristotle: A Reading of EE 5.12/ NE 6.12', *Journal of the History of Philosophy*.
- . ms. 'Causality and Intensionality'.
- Plantinga, Alvin. 1993a. *Warrant and Proper Function* (Oxford University Press).
- . 1993b. *Warrant: The Current Debate* (Oxford University Press).
- Pritchard, Duncan. 2007. 'Anti-luck epistemology', *Synthese*, 158: 277-97.
- . 2008. 'Sensitivity, safety, and anti-luck epistemology.' in John Greco (ed.), *The Oxford Handbook of Skepticism* (Oxford University Press).
- . 2009. 'Safety-Based Epistemology: Whither Now?', *Journal of Philosophical Research*, 34: 33-45.
- Putnam, Hilary. 1975. 'The meaning of 'meaning'', *Minnesota Studies in the Philosophy of Science*, 7: 131-93.
- Radford, Colin. 1966. 'Knowledge: By Examples', *Analysis*, 27: 1-11.
- Railton, Peter. 2006. 'Normative Guidance.' in Russ Shafer-Landau (ed.), *Oxford Studies in Metaethics, Volume 1* (Oxford University Press).
- Raz, Joseph. 1981. 'The Authority of Law', *Mind*, 90: 441-43.
- . 1985a. 'Authority and justification', *Philosophy & Public Affairs*: 3-29.
- . 1985b. 'Authority, Law and Morality', *The Monist*, 68: 295-324.
- . 2010. *Between Authority and Interpretation: On the Theory of Law and Practical Reason* (Oxford University Press).
- Reid, Thomas, and Derek R Brookes. 2000. *An inquiry into the human mind on the principles of common sense: a critical edition* (Penn State Press).
- Rescher, Nicholas. 1969. *Introduction to Value Theory* (Englewood Cliffs, N.J., Prentice-Hall).
- Reynolds, Steven L. 2002. 'Testimony, knowledge, and epistemic goals', *Philosophical Studies*, 110: 139-61.
- Rödl, Sebastian. 2007. *Self-consciousness* (Harvard University Press).
- . 2011. 'Forms of practical knowledge and their unity.' in Anton Ford, Jennifer Hornsby and Frederick Stoutland (eds.), *Essays on Anscombe's Intention* (Harvard University Press).
- Ross, Angus. 1986. 'Why Do We Believe What We Are Told?', *Ratio*: 69-88.
- Rothstein, Susan. 2004. *Structuring events: A study in the semantics of lexical aspect* (Blackwell: Malden, MA).
- Rowe, William L. 1979. 'The Problem of Evil and Some Varieties of Atheism', *American Philosophical Quarterly*, 16: 335 - 41.

- Schechter, Joshua. 2017. 'No Need for Excuses: Against Knowledge-First Epistemology and the Knowledge Norm of Assertion.' in J. Adam Carter, Emma C. Gordon and Benjamin W. Jarvis (eds.), *Knowledge First: Approaches in Epistemology and Mind* (Oxford University Press: Oxford).
- Schwenkler, John. 2015. 'Understanding 'Practical Knowledge'', *Philosophers' Imprint*, 15.
- Searle, John R. 1983. *Intentionality: An Essay in the Philosophy of Mind* (Cambridge University Press).
- . 1984. *Minds, Brains and Science* (Harvard University Press: Cambridge, MA).
- Setiya, Kieran. 2007. *Reasons without rationalism* (Princeton University Press).
- . 2008. 'Practical knowledge', *Ethics*, 118: 388-409.
- . 2009. 'Practical Knowledge Revisited', *Ethics*, 120: 128-37.
- . 2011. 'Knowledge of intention.' in Anton Ford, Jennifer Hornsby and Frederick Stoutland (eds.), *Essays on Anscombe's Intention* (Harvard University Press).
- Shah, Nishi. 2003. 'How truth governs belief', *Philosophical Review*, 112: 447-82.
- . 2008. 'How Action Governs Intention', *Philosophers' Imprint*, 8: 1-19.
- Shah, Nishi, and David J. Velleman. 2005. 'Doxastic deliberation', *Philosophical Review*, 114: 497-534.
- Slote, Michael. 1979. 'Assertion and belief', *Papers on language and logic*: 177-90.
- Small, Will. 2012. 'Practical Knowledge and the Structure of Action', *Rethinking Epistemology*, 2: 133-228.
- Smith, Michael R. 1994. 'The moral problem'.
- Sosa, Ernest. 1999. 'How to defeat opposition to Moore', *Noûs*, 33: 141-53.
- . 2000. 'Skepticism and Contextualism', *Noûs*, 34: 1-18.
- . 2006. 'Knowledge: Instrumental and testimonial.' in Jennifer Lackey and Ernest Sosa (eds.), *The Epistemology of Testimony* (Oxford University Press).
- . 2007. *A Virtue Epistemology: Apt Belief and Reflective Knowledge, Volume I* (Oxford University Press).
- . 2010. 'How competence matters in epistemology', *Philosophical Perspectives*, 24: 465-75.
- . 2015. *Judgment and agency* (Oxford University Press, USA).
- Srinivasan, Amia. 2015. 'Are We Luminous?', *Philosophy and Phenomenological Research*, 90: 294-319.
- . ms. 'Radical Externalism'.
- Stanley, Jason. 2011. *Know How* (Oxford University Press).
- Stanley, Jason, and Timothy Williamson. 2016. 'Skill', *Noûs*.
- Stathopoulos, Alexander. 2016. 'Knowing Achievements', *Philosophy*, 91: 361-74.
- Steward, Helen. 1997. *The Ontology of Mind: Events, Processes, and States* (Oxford University Press).
- . 2012. 'Actions as processes', *Philosophical Perspectives*, 26: 373-88.
- Stout, Rowland. 1997. 'Processes', *Philosophy*, 72: 19-27.
- Stroud, Sarah. 2006. 'Epistemic partiality in friendship', *Ethics*, 116: 498-524.
- Szabó, Zoltán Gendler. 2008. 'Things in progress', *Philosophical Perspectives*, 22: 499-525.
- Thomason, Richmond H. 1990. 'Accommodation, Meaning, and Implicature: Interdisciplinary Foundations for Pragmatics.' in Jerry Morgan Philip Cohen, and Martha Pollack (ed.), *Intentions in communication* (MIT Press: Cambridge, MA).

- Thompson, Michael. 2004. 'What is it to wrong someone? A puzzle about justice.' in R. Jay Wallace, P. Pettit, S. Scheffler and M. Smith (eds.), *Reason and Value: Themes from the Moral Philosophy of Joseph Raz* (Clarendon Press).
- . 2008. *Life and action* (Harvard University Press).
- . 2011. 'Anscombe's Intention and practical knowledge.' in Anton Ford, Jennifer Hornsby and Frederick Stoutland (eds.), *Essays on Anscombe's Intention* (Harvard University Press).
- Turri, John. 2016. 'The point of assertion is to transmit knowledge', *Analysis*, 76: 130-36.
- Unger, Peter. 1975. *Ignorance: A Case for Scepticism* (Oxford University Press USA).
- Velleman, David. 1989. *Practical Reflection* (Princeton University Press).
- Velleman, J David. 2001. "Practical Reflection " In. Stanford, CA: David Hume Series: CSLI Publications.
- Vendler, Zeno. 1957. 'Verbs and times', *Philosophical Review*, 66: 143-60.
- Vogel, Jonathan. 2010. 'Luminosity and Indiscriminability', *Philosophical Perspectives*, 24: 547-72.
- Watson, Gary. 2004. 'Asserting and promising', *Philosophical Studies*, 117: 57-77.
- Weatherson, Brian. 2004. 'Luminous Margins', *Australasian Journal of Philosophy*, 82: 373-83.
- Wedgwood, Ralph. 2002. 'The aim of belief', *Philosophical Perspectives*, 36: 267-97.
- . 2013a. 'Doxastic Correctness', *Aristotelian Society Supplementary Volume*, 87: 217-34.
- . 2013b. 'II—Doxastic Correctness', *Aristotelian Society Supplementary Volume*, 87: 217-34.
- . 2013c. 'The Right Thing to Believe.' in Timothy Chan (ed.), *The Aim of Belief* (Oxford University Press).
- Weiner, Matthew. 2005. 'Must we know what we say?', *The philosophical review*, 114: 227-51.
- . ms. "The Assurance View of Testimony." In.
- Williamson, Timothy. 2000. *Knowledge and its Limits* (Oxford University Press).
- . 2005. 'Replies to critics', *Philosophy and Phenomenological Research*, 70.
- . 2015. 'Justifications, Excuses, and Sceptical Scenarios.' in Julien Dutant and Fabian Dorsch (ed.), *The New Evil Demon* (Oxford University Press: Oxford).
- . forthcoming. 'Acting on Knowledge. .' in JA Carter, E. Gordon and B. Jarvis (eds.), *Knowledge-First* (Oxford University Press: Oxford).
- Witek, Maciej. 2013. 'How to Establish Authority with Words: Imperative Utterances and Presupposition Accommodation.' in Anna Brożek (ed.), *Logic, Methodology and Philosophy of Science at Warsaw University, Warszawa 2013*.
- Wolfson, Ben. 2012. 'Agential Knowledge, Action and Process', *Theoria*, 78: 326-57.
- Wright, Crispin. 2014a. 'Comment on Paul Boghossian, "What is inference"', *Philosophical Studies*, 169: 27-37.
- Wright, Stephen. 2014b. 'Sosa on knowledge from testimony', *Analysis*, 74: 249-54.
- Yaffe, Gideon. 1995. 'Velleman on Intentions as Reasons for Action', *Analysis*, 55: 107 - 15.
- . 2010. *Attempts: In the Philosophy of Action and the Criminal Law* (Oxford University Press).

- Zagzebski, Linda. 1994. 'The inescapability of Gettier problems', *Philosophical Quarterly*, 44: 65-73.
- Zagzebski, Linda Trinkaus. 2015. *Epistemic authority: A theory of trust, authority, and autonomy in belief* (Oxford University Press).