

Excerpts from the pre-publication version—please cite/quote from the published version.

Self-Reflection for the Opaque Mind

An Essay in Neo-Sellarsian Philosophy

T. Parent

SELF-REFLECTION FOR THE OPAQUE MIND attempts to solve a grave problem about critical self-reflection. The worry is that we critical thinkers are all in “epistemic bad faith” in light of what psychology tells us. After all, the research shows not merely that we are bad at detecting “ego-threatening” thoughts *à la* Freud. It also indicates that we are ignorant of even our *ordinary* thoughts—e.g., reasons for our moral judgments of others (Haidt 2001), and even mundane reasons for buying one pair of stockings over another! (Nisbett & Wilson 1977) However, reflection on one’s thoughts requires knowing what those thoughts are in the first place. So if ignorance is the norm, *why attempt self-reflection?* The activity would just display *naivety about psychology*. Yet while respecting all the data, this book argues that, remarkably, we are sometimes *infallible* in our self-discerning judgments. Even so, infallibility does not imply indubitability, and there is no Cartesian ambition to provide a “foundation” for empirical knowledge. The point is rather to explain how self-reflection as a rational activity is possible.

*To my mother and father,
those great Catholic educators*

Contents

Part I: Preliminaries

- Preamble: Is Philosophy Anti-Scientific?
- 1. Introduction: How is Rational Self-Reflection Possible?
- 2. The Empirical Case against Infallibility

Part II: Knowledge of Thought

- 3. Infallibility in Knowing What One Thinks
- 4. Objection 1: It's Apriori that Water Exists
- 5. Objection 2: Thought Switching
- 6. Content Externalism Does Not Imply Wayward Reflection

Part III: Knowledge of Judging

- 7. Infallibility in Knowing What One Judges
- 8. Infallibility in Knowing What One Expresses
- 9. Objection 1: It's Apriori that the Mental Exists
- 10. Objection 2: Attitude Switching
- 11. Attitude Confabulation Does Not Imply Wayward Reflection

Part IV: Denouement

- 12. Conclusion: How Rational Self-Reflection is Possible

How to Read This Book: The book aspires to be as user-friendly as possible. For this reason, some technical sections are marked with an asterisk '*'. These portions are intended mostly for specialists, and can be skipped by the more casual reader without losing the main thread of the book. (Ditto for the footnotes and appendixes.)

Detailed Table of Contents

<i>Chapter Abstracts</i>	
<i>Preface</i>	

Part I: Preliminaries

Preamble: Is Philosophy Anti-Scientific?	
0.1 A Neo-Sellarsian Approach	
0.1.1 Reconciling the images.....	
0.1.2 Against unconditional unity	
0.2 The Manifest Image and Quinean Scruples	
0.3 How is an Inescapable Fiction Possible?	
0.4 Transitional.....	
0.5 The “Categorical Interrogative”	
0.6 Commitment to Belief and to Self-Knowledge.....	
0.7 Some Clarifications	
0.7.1 The Kantian gambit.....	
*0.7.2 Is folk psychology presupposed?.....	
0.8 Closing Remarks	
Appendix: Philosophy as of Categorical Value	
References	

1. Introduction: How is Rational Self-Reflection Possible?	
1.1 The Problem of Wayward Reflection.....	
1.2 Methodological Crisis in Philosophy and Psychology.....	
1.3 The Start of an Answer.....	
1.3.1 Opening objections.....	
*1.3.2 Internal- and external-world skepticisms	
1.3.3 Epistemic internalism?	
1.4 Overview of the Book.	
References.	

2. The Empirical Case against Infallibility	
2.1 What is Evidence-Based Antagonism?	
2.2 Revisiting the Experiments	
2.3 The Argument against Evidence-Based Antagonism.....	
2.4 Objections and Replies.....	
*2.4.1 Metacognitive anti-representationalism	
*2.4.2 The objection from the intentional stance	
*2.4.3 The rejoinder from expressive speech acts.....	
2.4.4 So what?	
2.5 Onward to Part III	
Appendix: Argument for the Entailment.....	
References	

Part II: Knowledge of Thought

3. Infallibility in Knowing What One Thinks	
3.1 Ground Clearing	
3.2 Burge's Way of Self-Knowing	
3.3 Infallibilism Derived	
*3.4 Necessity and the Apriori	
3.5 Looking Ahead	
References	
4. Objection 1: It's Apriori that Water Exists	
4.1 The Objection	
4.2 Externalism about Empty Concepts	
*4.3 A Problem with Form?	
*4.4 Concept Possession	
4.5 The "Revenge" Argument	
4.6 Transition to Chapter 5	
References	
5. Objection 2: Thought Switching	
5.1 The Objection	
5.2 Introspective Awareness and Skeptic-Immunity	
5.3 Knowing-That versus Knowing-What	
5.3.1 Knowing-what and informativeness	
*5.3.2 Supplementary remarks	
5.4 Purpose Relativity and Knowing-What	
5.5 Transition to Chapter 6	
Appendix: Quasi-Intellectualism about Knowing- <i>Wh</i>	
References	
6. Content Externalism Does Not Imply Wayward Reflection	
6.1 Boghossian on Wayward Reflection	
*6.2 The Owens-Brown Argument	
*6.3 Non-Introspectable Equivocation <i>sans</i> Externalism	
*6.4 Internalist Replies	
6.5 Tacit Belief and Identity	
6.6 Further Dialectics	
*6.6.1 Vicious regress?	
*6.6.2 Collins' objection and related matters	
6.6.3 Explicit equivocation	
6.7 Onward to Part III	
References	

Part III: Knowledge of Judging

7. Infallibility in Knowing What One Judges	
7.1 Chapter 3 Redux	
7.2 Paratactic Judging?	
7.3 Lagadonian Reference	
7.4 Semi-Lagadonian Judging	
*7.5 Infallibility without Mentalish	
7.6 Link with the Next Chapter	
Appendix: Infallibility as an Ideal of Rationality	
References	
8. Infallibility in Knowing What One Expresses	
8.1 Introduction to the Criterion	
8.2 Some Terminology	
8.3 The Criterion Again and a Sketch of Its Argument	
8.4 Evansian Autological Examination	
8.4.1 Autoexams and replies to autoexams	
*8.4.2 Some complications	
8.5 Argument for the Criterion	
8.5.1 Sincerity as criterial	
8.5.2 Reflex-like asserting implies sincerity	
8.4.3 Review, plus the criterion extended	
8.6 Recognizing the Sign	
8.7 Anti-Intellectualism about Expressive Know-How	
*8.8 Reflex-like Lying	
8.9 Foreshadowing	
Appendix: Expressive Showing Enables Semi-Lagadonian Judging	
References	
9. Objection 1: It's Apriori that the Mental Exists	
9.1 The Objection	
9.2 What is Mental Fictionalism?	
9.3 Worries about Self-Refutation	
*9.4 Semantic Ascent	
*9.5 Against Primitivism	
9.6 Mental Fictionalism <i>cum</i> Quietism	
9.7 Transition to Chapter 10	
References	
10. Objection 2: Attitude Switching	
10.1 The Objection	
10.2 Awareness of Semi-Lagadonian Judging	
10.3 Purpose Relativity and Evansian Autoexams	
*10.4 Compare and Contrast	
10.5 Transition to Chapter 11	
References	

11. Attitude Confabulation Does Not Imply Wayward Reflection.....

 11.1 Wayward Reflection and the Attitudes

 11.2 Re-appropriating *qua* Evidence.....

 *11.3 Response to Kornblith.....

 11.4 Reflection Demands Expression

 References

Part IV: Denouement

12. Conclusion: How Rational Self-Reflection is Possible

 12.1 Self-Reflection for the Opaque Mind.....

 12.2 Neo-Sellarsian Farewell

 References

Index.....

Chapter Abstracts

Preamble: *Is Philosophy Anti-Scientific?* Science often conflicts with our everyday experience. The book is an examination of such a conflict: Psychological studies show that we are unreliable in knowing our own beliefs and desires—yet from the everyday point-of-view, self-reflection seems like a rational activity. But how can it be rational, since it takes for granted knowledge of one’s own mind? For Sellars, philosophy’s aim is to resolve such discrepancies between the “manifest” and “scientific” images. However, some will protest that philosophers should not “negotiate” with science—the scientific image should instead claim hegemony. However, I defend the Sellarsian by arguing that we are simply unable to jettison central parts of the manifest image. That is so, even if these parts are fictitious. If we are stuck with them, moreover, then the Sellarsian cannot be blamed for trying to reconcile them with science, even if this is less than scientifically ideal. The “inescapability” thesis is argued by suggesting that the question ‘What should I do?’ is “forced” in roughly William James’ sense. Very briefly, whatever you do rationally commits you to your act being what you should do—which in turn, assumes the existence of norms. So in that sense norms are “inescapable.” The argument is then extended to show that commitments to belief and to some kinds of self-knowledge are similarly inescapable.

1. ***Introduction: How is Rational Self-Reflection Possible?*** This chapter introduces the main project of the book—to show that self-reflection can be a rational activity, despite the psychological considerations for widespread self-ignorance. (This is dubbed the “Problem of Wayward Reflection.”) It is also illustrated how the Problem bears importantly on both philosophical and psychological methodology.
2. ***The Empirical Case against Infallibility.*** Philosophers and psychologists generally hold that, in light of the empirical data, a subject lacks infallible access to her own mental states. However, while subjects certainly are fallible in some ways, this chapter shows that the data fails to discredit that a subject has infallible access to her own *occurrent thoughts and judgments*. This is argued, first, by revisiting the empirical studies, and carefully scrutinizing what is shown exactly. Second, I argue that if the data were interpreted to rule out all such infallibility, the relevant psychological studies would be self-effacing. For they adopt a methodology where a subject is simply presumed to know her own second-order thoughts and judgments—as if she were infallible about them. After all, what she expresses as her second-order judgment is trusted as accurate without independent evidence, even though such judgments often misrepresent the subject’s first-order states. The upshot is that such studies do not discredit all infallibility hypotheses regarding one’s own mental states.
3. ***Infallibility in Knowing What One Thinks.*** This chapter defends the view that a subject has infallible judgments about the contents of her occurrent thoughts. First, a contemporary defense of this claim is examined, given by Burge (1988), yet it is found lacking. I then offer my own defense, appealing to a minimal version of the language of thought hypothesis. The argument here has the virtue of refraining from any semantic premises; thus, it is congenial to both internalists and externalists about content. Still, the argument is not itself regarded as infallible, given that some of the premises are empirical. For that reason, the present infallibilism does not preclude the Quinean thesis that every belief is vulnerable to empirical revision. Even so, the argument also reveals how a subject can know her own thoughts while in the armchair.

- 4. *Objection 1: It's Apriori that Water Exists.*** Several authors have objected that, assuming we have armchair knowledge of our own thoughts, content externalism implies that we can armchair-know contingent facts about the empirical world. After presenting the objection, I reply by resisting the premise that an externalist can armchair-know: If s/he has the concept water, then water exists. In particular, Boghossian's Dry Earth example suggests that such thought-experiments do not provide this kind of armchair knowledge. Boghossian himself rejects the Dry Earth experiment, however, since it would imply that externalism is true of empty concepts as well as non-empty concepts. Yet in this chapter I respond by defending empty-concept externalism, from criticisms by Boghossian and Brown, and developed further by Besson. My contention is that an externalist can give a non-ad hoc descriptivist account of empty concepts. Accordingly, armchair self-knowledge does not enable an externalist to know contingent features of the external world.
- 5. *Objection 2: Thought Switching.*** In this chapter, a different objection is examined against externalist armchair knowledge of thought-content: the slow-switch argument. Following Goldberg, the argument is construed a challenge to the externalist's *skeptic immune* self-knowledge (where such knowledge would persist under various skeptical hypotheses). Goldberg attempts to reclaim such knowledge for the externalist by developing a strategy from Burge. Nevertheless, it is noted that such Burge-style accounts only address a subject's ability to *know that* she is thinking that "water is wet." They do not explicitly concern the subject's ability to *know what* she is thinking, which is the distinctive type of knowing at issue in the slow-switch argument. Subsequently, however, the Burge-style view is recast so that the relevant "knowing what" has a chance against the argument. For one, it is emphasized that such "knowing what" should be construed as *ontologically non-committal* (so that knowing what a water-thought is does not require knowing that water exists). Second, following Boër & Lycan, "knowing what" is construed as *purpose relative*, in that whether one "knows what" depends on whether one knows enough to achieve some contextually salient purpose/goal. And for at least some purposes, it is defensible that an externalist can introspectively "know what" she thinks—even in a skeptical context.
- 6. *Content Externalism Does Not Imply Wayward Reflection.*** This chapter opposes Boghossian's argument that externalist armchair self-knowledge leads to Wayward Reflection. Boghossian's charge is that such knowledge does not prevent "non-introspectable equivocation," i.e., equivocation that cannot be detected introspectively. Boghossian says such equivocation would be possible under externalism, since externalism implies that the content of 'water' (and of the concept it expresses) might change from premise to premise, owing to stealthy "slow switches" between twin environments. Moreover, because the change owes to environmental differences, the resulting equivocation could not be known introspectively. In reply I argue two points. First, I suggest that if non-introspectable equivocation is a problem, it is a problem that is not unique to externalism. Second, I show that non-introspectable equivocation is ultimately not a problem. That's because (barring gross psychological deviance) a subject will assume that a concept and a twin-concept are co-referring. Consequently, the reasoning is normally not a case of equivocation, but rather one where a false heterophonic identity-statement is assumed. (Others have said that this leads to a vicious regress; however, I show that the regress is in fact a virtuous one.) And though this still counts as an error, it is not the sort of error that forces a question about the basic rationality of the subject.

- 7. *Infallibility in Knowing What One Judges.*** This chapter defends a strong infallibilism about knowing what one *judges*. The claim is that in a specified set of cases, “I am judging that I am judging that *p*” semantically entails “I am judging that *p*.” The chapter begins by reviewing the weaker, content-infallibilism from chapter 3; this serves to introduce nine caveats which are carried over to the infallibility thesis about judging. Paratactic second-order judgments are then introduced; these are characteristically of the form “I am judging *that: p*,” where ‘*that*’ is a deictic term denoting an act of first-order judging. But though this suggests a kind of “semantic locking” between the first- and second-order judgments, the possibility of empty deixis means that strict infallibility is not yet secured. In response, the infallibility thesis is restricted to cases of judging where deictic reference is replaced by “Lagadonian reference.” (The term ‘Lagadonian’ is from Lewis 1986, borrowing from Jonathan Swift.) Lagadonian reference is where an object is used as a name for itself; several examples from ordinary linguistic practice are provided. The hypothesis, then, is that an act of first-order judging can also function as a Lagadonian name for itself—and when this judging-act is self-attributed by means of its Lagadonian name, the reference is inevitably non-empty, and the self-attribution is strictly infallible.
- 8. *Infallibility in Knowing What One Expresses.*** Here, it is shown that in linguistic practice, there is a “sure sign” for when a speaker executes an infallible second-order judgment. Specifically, an unplanned and automatic (“reflex-like”) assertion of “*p*” inevitably co-occurs with a true judgment that *I judge that p*, when the assertion acts as a possible reply to one type of question. This is called an “autological examination;” it is paradigmatically the type of question asked by a pollster. Evans’ query ‘Will there be a World War III?’ is considered as an example (although different readings of the question are distinguished). The infallibility thesis is shown, first, by defending a certain “dual expression” thesis: If a reflex-like reply to an Evansian autoexam like ‘Will there be a WWIII’ is ‘Yes’ (where this is uttered in a competent way), the reply expresses *both* the speaker’s first-order judgment that there will be a WWIII, and simultaneously, a self-attribution of that self-same judgment. This “dual expression” thesis then leads immediately to the infallibility claim. The second-order judgment, expressed in a reflex-like reply to an Evansian autoexam, is invariably true in virtue of the co-expressed first-order judgment. (The occurrence of the first-order judgment is precisely what the second-order judgment contends.) Further, in the interest of preventing Wayward Reflection, the chapter explains how one might go about *recognizing* the “sure sign.” Utterances may be assumed competent by default (cf. Wright 1998; 2000), where competence is a type of (non-intellectualist) know-how, acquired by conditioning in the linguistic community. The reflex-like feature evidences that an uttering is *sincere*, and is recognizable by a quick response-time, its semi-involuntary occurrence, etc. (An utterance’s status as a *reply* to a given autoexam is a matter of knowing the *content* expressed—and on such knowledge, the reader is referred back to Part II.) Finally, the chapter ends by sketching how all this will bear in later chapters on the Problem of Wayward Reflection.
- 9. *Objection 1: It’s Apriori that the Mental Exists.*** Armchair self-knowledge would seem to enable armchair knowledge that eliminativism is false. For if I armchair-know *what* I believe, then I can armchair-know *that* I have beliefs. In reply, this chapter argues that the possibility of *mental fictionalism* blocks this worry. Mental fictionalism holds that the mental states posited by folk psychology do not exist, yet that some sentences of folk psychological discourse are nonetheless “true,” when understood as “truths according to the mentalistic fiction.” But after formulating

this view, I identify five ways that it appears self-refuting. More, it is shown that self-refutation cannot be avoided either by semantic ascent or by a kind of primitivism. Even so, I eventually show that in speaking of “self-refutation,” objection is subtly question-begging. So in the end, the possibility of mental fictionalism remains standing, and the armchair refutation of eliminativism is thereby subverted.

- 10. *Objection 2: Attitude Switching.*** This chapter considers objection that, since one cannot discriminate from the armchair discriminate an attitude (like judging) from all relevant alternative attitudes, one cannot armchair-know what one’s attitude is toward a proposition. This is buttressed by the actual psychological evidence on so-called “moral dumbfounding.” In reply, I argue that knowing what attitude one has might not always require the discriminatory ability in question, thanks to the purpose-relativity of “knowing what” (cf. chapter 5). Indeed, it is defensible that one knows what one’s attitude is, for the purpose of appropriately replying to an Evansian autoexam. Or at least, that is tenable, assuming one’s reply expresses a semi-Lagadonian judgment (cf. chapter 7), and the reply is seen as criterial for such a judgment (cf. chapter 8). But unlike chapter 5, I do *not* defend the pertinent “knowing what” in the context of an external-world skeptic. Nonetheless, it is contended that one can “know what” for Evansian autoexam purposes, even in a context where data on moral dumbfounding and the like are salient (a.k.a., an “antagonistic” context; cf. chapter 2). This also explains why experimental psychology is right to trust subjects in expressing their own judgments, even though such subjects are revealed to be largely self-ignorant. The account of self-knowledge is also contrasted with other accounts in the contemporary literature.
- 11. *Attitude Confabulation Does Not Imply Wayward Reflection.*** Suppose, per the last chapter, that one “knows what” attitude one has toward p , for the purposes of appropriately replying to an Evansian autoexam. Still, it does not follow that one “knows what” for the purposes of avoiding Wayward Reflection. This is illustrated by a subject who is unable to armchair-discriminate her stably affirming p versus affirming p while also having a settled disposition to deny p . In response, I argue that if a subject’s judgments are expressed in writing, she can gain evidence that those judgments are not merely temporary aberrations, by exercising—and bearing witness to—her acts of “re-appropriating” graphically expressed judgments. (Re-appropriating is shown to be highly analogous to asserting as per chapter 8, since both types of act normally occur with the same type of communicative intent.) In brief, re-appropriating provides evidence of a judgment’s “staying power,” thus providing some reassurance against Waywardness. Finally, such considerations are used to counter Kornblith’s antagonistic arguments on self-reflection.
- 12. *Conclusion: How Rational Self-Reflection is Possible.*** This is a short, concluding chapter sketching the overall picture yielded by what precedes it. Besides summing up how self-reflection can be a rational activity, it also offers a few final comments on Neo-Sellarsian metaphilosophy.

Preface

Why a book on reflection? One reason is: Timing. These days in philosophy, there is a frenzy of interest in metaphilosophical issues, and *philosophical method* looms large in this. Reflection, moreover, is the paradigmatic philosophical method. (The book was almost entitled “Philosophical Method for the Opaque Mind.”) Another reason: A gap. The book presents an important problem about reflection which has been mostly ignored in the literature, and is currently the sole attempt to resolve it.¹ Third: Broad appeal. The issue about “reflection” is at least a problem about reflective *critical thinking*. And the problem does not just consist in some esoteric philosophical riddle. Rather, it arises straightforwardly from concrete empirical results in psychology. I refer to evidence suggesting that we are *largely ignorant of our own minds*. The problem is that if we are largely self-ignorant, then self-reflection (e.g., critical scrutiny of one’s own thoughts) makes no rational sense. One would be ignorant of the thoughts one is supposed to reflect on.

The book thus was written to defend the rationality of self-reflection despite the evidence for widespread self-ignorance. In this, the book also purports to show by example how philosophical work on a *prima facie* empirical topic can be not only appropriate but edifying as well. (Cf. the Preamble chapter.)

¹ In fact, one version of the problem has been discussed extensively; I am referring to the literature spawned by Boghossian’s “inference argument” against content externalism. Yet the more daunting version—examined in chapter 11—is gestured at only in the work of Kornblith. (It is the more daunting version since a full solution, I think, requires a solution to Boghossian’s version.) Regardless, unlike myself, Kornblith largely capitulates to the difficulty. (For pertinent works by Boghossian and Kornblith, see the references to chapter 1.)

Throughout, there is a preoccupation with philosophical method, not only in theory, but also in the book's own practice. The book attempts to meet certain standards of "rigor," which makes the work more difficult in some respects. But there is a trade-off I think that makes this necessary. Allow me to explain.

What does the theory of mind (hereafter, "TM") hope to achieve? Uncontroversially, it is to advance our understanding of the mind. For my part, this requires developing TM into an increasingly *scientific* theory—but what does that mean? The immediate thought is to align TM more and more with empirical data. However, empirical data can be brought to bear on TM only if it is *already* scientific in certain ways. Most obviously, TM must already have definite, observable consequences that can be tested against the data. But to a significant extent, TM has not yet reached that point. The question, then, is how do we get from here to there?

A mature TM, restricted to human minds, can be seen as a branch of human anatomy. (This is not to endorse reductionism; it is only a point about the relation between disciplines.) It's not just that some brain anatomy is necessary for understanding our minds. More basically, it is that our mental state-tokens are token-states of the body (at least in part). And so, a complete inventory of bodily states, per human anatomy, will include an inventory of mental states (or the bodily components thereof).

However, a salient feature of anatomy is its copious taxonomy. Even a relatively simple body-part, such as the forearm, has been mapped out in abundant detail. For starters, there are the two bones, the radius and the ulna, and the interosseous membrane connecting them. There are also numerous muscles, including the brachioradialis, the pronators, and supinators—plus the veins permeating the area, the main ones being the cephalic, median, antebrachial, and basilic veins. And this is not yet to mention the nerves. Already, we have a good amount of technical

terms. But here, it is obvious that precision in language enables a more refined understanding; it is what permits (e.g.) a surgeon to focus an operation on “the annular ligament of the ulna.”

These days, however, some theorists of mind proceed as if precise language is optional. Yet how much more subtle is the anatomy of the mental! If the aim is the scientific development of TM, then, we must not shy from precise terminology. Again, exact language is often what allows experts to make distinctions which novices cannot. I apologize, then, if my prose is not as readable as some TM-writers. Yet from where I sit, it is no longer excusable to write as if causal uses of ‘representation’, ‘content’, or ‘belief’ are sufficiently clear for TM.² Part of what makes a theory empirically testable is that its predictions have a kind of *determinacy* to what they say. This helps guarantee that “accommodating the data” is not something the theory can fudge, by exploiting the vagueness of its language. That means imposing precision. Precise terminology may make the work of TM more difficult, but it makes it better. (Yet to aid the reader, the definition of key terms are repeated where relevant; this also helps make each chapter readable independent of the others.)

Having said all that, a *wealth* of specialized terminology in TM does not yet make sense. Thus the book is not as lexically adventurous as it might have been, and it is not yet primed for serious empirical testing. Ultimately the language is more meticulous than much philosophy of mind, but I hope to have explained why this heads in the right direction.

I have incurred a number of debts to individuals and institutions in the course of writing. I am most grateful to my family, especially to BZ and to my parents, for their uplifting presence and therefore positive influence on the writing process. I am also particularly grateful to two

² It is permissible to *start* by using such terms in the ordinary way, as is true in the present work. But my point is that the ordinary use is not where TM should end.

colleagues: Ram Neta and Joseph C. Pitt. Although they may not realize it, their enthusiasm was tremendously encouraging at a professionally difficult time. I also must give special recognition to Dorit Bar-On, William G. Lycan, and (the late) Jay F. Rosenberg. Their philosophical influence on this work is unmistakable, and I thank them for all the erudition over the years. Also, special appreciation goes to Margo Irvin and Andrew Weckenmann at Routledge for their support of the project.

My gratitude is sincerely extended to other individuals for feedback and/or discussion on issues in this book, including Louise Antony, Jody Azzouni, Jim Baillie, Derek Ball, Gary Bartlett, Jason Bowers, Tyler Burge, Yuri Cath, David Chalmers, Matthew Chrisman, Larry Coulter, Chris Daly, Felipe De Brigard, Justin Fisher, Tamra Frei, Matthew Fulkerson, Brie Gertler, Ephriam Glick, Sandy Goldberg, Glen Hoffmann, Thomas Hofweber, Bryce Huebner, Ben Jantzen, Justin Jeffrey, Jim Klagge, David Landy, Douglas Lind, Dan Linford, Joe Levine, Doug Long, Eric Mandelbaum, Norah Martin, Aidan McGlynn, Tristram McPherson, Ruth Millikan, Greg Novack, Nico Orlandi, Mark Phelan, John Post, Jesse Prinz, Amber Ross, Dan Ryder, Sarah Sawyer, Keith Simmons, Kelly Trogon, Meg Wallace, and Brian Weatherson. (I am certainly forgetting some names; this has been ongoing for years. My deepest apologies.)

The contents of the book have also benefited from audiences at meetings of the American Philosophical Association (2010 Central Division, 2004 Eastern Division), the University of Edinburgh (July 2014 conference on Mental Fictionalism), the University of Glasgow (June 2013 conference on the philosophy of Tyler Burge), Oregon State University (2012 Northwest philosophy conference), the University of Memphis (2004 Midsouth philosophy conference), the University of North Carolina at Chapel Hill (talk in Oct. 2008), Vanderbilt University (talk in Oct. 2009), and Virginia Tech (April 2015 workshop in the philosophy of mind).

Also, my thanks to the publishers below for the use of material previously in print:

- Chapter 2 is a revised/expanded version of “The Empirical Case against Infallibilism,” *Review of Philosophy and Psychology*, 7.1 (2016): 223-242.
- Chapter 3 is mostly a revised/expanded version of “Infallibilism about Self-Knowledge,” *Philosophical Studies*, 133.3 (2007): 411–424. The exception is section 3.4, which is a revised/expanded version of “Infallibility Naturalized: Reply to Hoffmann,” *Dialectica*, 67.3 (2013): 353-358.
- Chapter 4 is a revised/expanded version of “Self-knowledge and Externalism about Empty Concepts,” *Analytic Philosophy*, 56.5 (2015): 158-168.
- Chapter 5 is a revised/expanded version of “Externalism and ‘Knowing What’ One Thinks,” *Synthese*, 192.5 (2015): 1337-1350.
- Chapter 9 is a revised/expanded version of “In the Mental Fiction, Mental Fictionalism is Fictitious,” *The Monist*, 96.4 (2013): 608-624.

The assertive mode of speech is needed for describing (and describing is what the book hopes to do). Yet asserting has the unfortunate feature of “representing the speaker as knowing,” as they say, even though what follows is highly tentative. But, rather than incessantly writing “it only *seems* to me that *p*,” I instead hereby cancel any implicature for knowing the propositions asserted.³ Hopefully I can rest easier.

TSP

Blacksburg, VA

³ Though per chapter 8, I do not cancel implicatures to knowing *that I judge* the propositions asserted. (That would create a formidable preface-paradox indeed.)

Preamble

Is Philosophy Anti-Scientific?

Initially, this book may seem to be philosophical excess of the worst kind. It can look like wishful thinking, a panicky response to legitimate scientific results about the limits of self-knowledge. For, despite ample evidence of these limits, the book argues that many beliefs about one's own mind are not just reliable; they can be *infallible* in the way Descartes suggested. Such a view, if not anti-scientific, will at least seem quixotic.

I shudder that this might be the impression left on the casual reader. Like many contemporaries, I think of myself as a naturalist—I not only reject ghosts and gods, but more generally accept that science is the highest authority on what mind-independent reality is like. (Yet this is not to say that science is perfect.) Thus, if science reveals that the mind is largely opaque to itself, one would expect a naturalist philosopher to let the matter lie.

However, nowhere is any of the scientific data contested. Quite the contrary, the data puts an inviolable constraint on the project. Still, even if the data shows the mind is *often* self-ignorant, it does not follow that this is *most often* the case. It at least remains possible that in many *important* cases, we are attuned to our own minds. Naturally, to suggest the possibility is not to demonstrate its actuality. Yet this illustrates that some first-person authority may be possible, even given the science. (Though whether it is at all probable is another matter.)

0.1 A Neo-Sellarsian Approach

Be that as it may, why exactly should this position be developed? (There is still the smell of wishful thinking.) The answer is found in the next chapter. As a prior matter, however: Why is

a *philosopher* suited to investigate the issue? Whether homo sapiens know their own minds may seem better left to empirical psychology, especially by the lights of a self-proclaimed naturalist.

0.1.1 Reconciling the images

I certainly welcome confederates from psychology. Nevertheless, some issues in the area are distinctly philosophical, in that they exist as conflicts between what Sellars (1962) called the *manifest image* and the *scientific image* of the world. The leading question of the book is an example: How is rational self-reflection possible? From the manifest perspective, self-reflection stands as an important part of our intellectual and moral lives. Yet the activity presumes a reliable ability to identify one's own beliefs, desires, etc., in order to evaluate them. And *reliable* access of this type appears non-existent in light of contemporary psychology.¹

So is self-knowledge and self-reflection mere foolishness? Either way, the point is that we have now entered distinctly philosophical terrain. The question is: What should we do about this conflict between the images? Should we try to resolve it? Or should we reject self-knowledge, and with it the practice of self-reflection? (Alternatively, should we just learn to live with the conflict?) Importantly, since the issue is a *normative* one, it seems appropriate to address it via (empirically informed) philosophy.

The book adopts a particular answer to the normative question—namely, that *resolving* the conflict is best. But this is contentious in two ways. First, it is doubtful whether the philosopher is qualified to discuss what exactly the data allows. Second, it may appear more

¹Also, self-knowledge is central to the manifest image since, apparently, it is widely presupposed in ordinary conversation. We assume by default that a speaker *knows what she is saying*. (Without that assumption, interpreting the speaker is rather more difficult.)

scientifically serious just to reject what the manifest image says, rather than prop it up somehow, by massaging away the empirical difficulties.

Regarding the first issue, a philosopher may well be unqualified to analyze the data for prototypical scientific purposes. Still, a Sellarsian is not necessarily out of line if she offers hypotheses to show how reconciliation is possible. Such hypothesis-building is certainly not the special province of philosophers. Yet if the Sellarsian recommends resolving conflicts between the images, she thereby incurs a burden to answer the “how possible?” question. She becomes *dialectically obligated* to do some amount of hypothesis-building, given that some of her peers answer the normative question differently. Whether her hypotheses are scientifically viable should be judged on a case-by-case basis. (And if they are not, that is reason to judge them philosophically lacking as well.) But even if they are *de facto* ignored by scientists, they can still serve a dialectical function vis-à-vis the normative question, a question which is indeed appropriate for philosophers.

The book thus develops an increasingly nuanced hypothesis about self-knowledge, meant to vindicate the manifest value of reflection. But the second issue, again, is that it would seem more scientifically respectable just to dismiss self-reflection as misguided. In fact, this is an instance of a very general problem for Sellarsian metaphilosophy—namely, that the Sellarsian looks *anti-science* insofar as she advocates conserving the manifest image, despite its conflicts with the scientific image (hereafter, “MI” and “SI”). What follows is an attempt to respond to that objection.

Even so, I do not endorse in an unqualified way Sellars’ metaphilosophy. Yet to fend off the anti-science objection, the main thesis shall be that engagement with the manifest image is simply *non-optional*. Further, if we are stuck with MI, we cannot be blamed for continuing to

engage it, even if this is less than scientifically ideal. In which case, it is only understandable that we would try to reconcile the conflicts with SI. We could not be blamed for trying.²

[The remainder of the preamble has been omitted.]

² A suppressed premise here seems to be ‘rational ought’ implies ‘psychological can’. However, I would not assume that this holds universally. Though I think something like it holds in this instance.

Chapter 1

Introduction: How is Rational Self-Reflection Possible?

Back in Catholic elementary school, reflection was mandatory. Time was set aside for regular prayer (to reflect on our blessings and shortcomings), and we were taught that the Watchful Eye of God was always upon us. The latter encouraged a certain habit of self-monitoring, to gauge how one's thought and action appeared from His point of view. Thus: "Catholic guilt," the feeling that whatever good one *can* do, one *ought* to do. I do not necessarily recommend raising children this way.

Regardless, Catholics are right that reflection is an important part of our intellectual and moral lives. Philosophers are generally sympathetic to this. Even so, philosophers have rightly abandoned *Cartesian* reflective inquiry as hopeless. It is not as if apriori philosophical reflection could justify *science* (or empirical inquiry in general). If anything, science reveals how unjustified philosophy is. Nonetheless, many of us believe that critical reflection has some benefit.¹ It is typically thought that such reflection can expose skewed priorities, inconsistencies, non-sequiturs in reasoning, and so forth. The practice is prized in many academic disciplines, but it seems especially central to philosophy—and philosophy, I would argue, serves critical reflection particularly well.

¹ Still, there is empirical evidence against the benefits of self-reflective inquiry. See, e.g., Wilson & Schooler (1991), Wilson & Kraft (1993), Wilson & Hodges (1993), Wilson et al. (1993), Damasio (1994), Wilson & Dunn (2004), Epley & Gilovich (2005), and Halberstadt & Wilson (2008). Kornblith (1999, 2002, ch. 4; 2010, 2012; 2013, 2016) is one who appeals to such evidence to ask pointedly why reflection should be such a revered practice. But while these doubts are quite important, I am unable to address them in the present work. Yet see Tiberius (2008; 2013) for a response to such doubts.

1.1 *The Problem of Wayward Reflection*

Nevertheless, anyone who champions self-reflection² faces a grave difficulty. The activity presumes epistemic access to one's own thoughts: I cannot critically reflect on my thoughts in ignorance of what those thoughts are. But unfortunately, there are strong reasons to say that each of us *are* largely ignorant of our own thoughts. For one, many argue that content externalism (sometimes called anti-individualism³) precludes knowing what one thinks just while seated “in the armchair” or without empirical investigation.⁴ Also, there is a second reason, less discussed in the philosophical literature. Experimental psychology indicates that we *are* ignorant, to a large degree, about our own thoughts—especially those thoughts that act as our *reasons* for action or

² How does reflection differ from *self*-reflection? I often use the two terms interchangeably, but strictly speaking, only the latter directly targets *one's own* beliefs. Thus, not all reflection is self-reflection proper—I can reflect on the beliefs of others, e.g., our political leaders. (Yet such reflection still operates on my beliefs about what the other person believes.) That may suggest, however, that self-reflection proper takes a *metacognitive* form, e.g., “I believe that social equality is just; therefore...” But not so: It can just feature my first-order beliefs: “Social equality is just; therefore...”. Still, it counts as *self*-reflection only if they are understood as *my* beliefs. (If I see them just as Obama's beliefs, there is a clear sense in which I am not *self*-reflecting, even if I happen to share his beliefs.)

³ See Putnam (1973; 1975), McGinn (1977), Stich (1978), Burge (1979; 1982; 1986). Burge prefers the term ‘anti-individualism’, yet that thesis is not really about *content* at all. (See Burge 2006.) In conversation, Burge clarifies that he holds Frege's view where contents exist as abstracta. Thus, anti-individualism more concerns individuating one's *concepts* rather than their contents. But Fregean contents are not assumed here; thus, I avoid ‘anti-individualism’, since the relevant non-internalist thesis is not really Burge's.

⁴ At times, the book vacillates between talk of “armchair,” “apriori,” and “introspective” self-knowledge. “Armchair” self-knowledge covers the other two, though it is not exhausted by them. Indeed, in Part III, I discuss a type of self-knowledge which, although attainable from the armchair, is somewhat empirical. (This is “expressive” self-knowledge; it is achievable in linguistically expressing one's judgments.)

belief.⁵ Yet such widespread self-ignorance would confound self-reflection. One would be ignorant of the thoughts one is supposed to reflect on.

These two points—re: externalism and experimental data—each motivate what I call “The Problem of Wayward Reflection” (cf. Parent 2007, p. 421). Confined to the armchair, how can we reliably discern our own thoughts, rather than just confabulating them? If the answer is “we can’t,” then armchair self-reflection makes no sense. If one cannot reliably know what one’s reasons *are*, then so much the worse for the hope to *evaluate* them.⁶

The basic Problem can be made vivid by an analogy. Suppose neuroscience has advanced to the point where we can use a “brain scanner” to detect what your thoughts are. When at peak performance, the scanner is even sensitive enough to detect thoughts that are otherwise non-conscious. Suppose, then, that for several months your therapist has used the scanner to identify a number of thoughts, and has focused your therapy toward acknowledging and evaluating them. But suppose it is discovered that the scanner has been malfunctioning for several weeks—and that the thoughts reported included a lot of false positives. Now although s/he is alarmed, suppose your therapist recommends continuing with this method anyway, as if nothing were wrong. Wouldn’t

⁵ For a list of pertinent works in experimental psychology, see n. 1 of chapter 2. N.B., I assume throughout that reasons for belief are themselves beliefs. (Yet I am neutral on whether reasons *qua* reasons are causes.) An objection: If I judge that the mail carrier has not arrived yet, my reason is not *that I have the belief* that the little flag on the mailbox is still raised. Rather, it is *the fact that* the flag is still raised. My preference, however, is to say that the latter is a piece of *evidence*, re: the mail carrier’s arrival. My access to this evidence still consists in a belief, and that is all I mean when speaking of my reasons as beliefs.

⁶ A Kantian ethicist should be doubly concerned; knowing her reasons is also needed to *morally* assess her own acts. (For the Kantian, the worth of her act depends on her “maxim.”) Ignorance here would thus force ignorance of, e.g., to what extent the “dear self” drives her act. (Further still, Lehrer 1990 argues that access to one’s beliefs is key to many phenomena of philosophical interest—including free action, rational acceptance, and even intentionality itself.)

you find this laughable? The thoughts dictated by the scanner might be interesting to think about independently, but they may be irrelevant to your therapeutic goals. You may end up dealing with thoughts that have no relevance to *your* aims and interests.

The Problem of Wayward Reflection suggests that we are actually in this dire relationship to our own mind. But the difference is that each person is dealing with a deficient *internal* scanner as it were. To repeat, externalist arguments suggest that armchair-methods cannot enable discriminations between one's thoughts and any number of "twins".⁷ This seems to imply that one might regularly and unwittingly *equivocate* during armchair reflection, due to a thought being "switched" with a twin (see Boghossian 1994a, b, 1994, 2010, 2014.). Call this "Wayward Reflection via Thought Switching." Further, studies show not just that we are bad at introspecting our own ego-threatening beliefs *à la* Freud, e.g., "My mother never loved me." The data also reveals that people are bad at introspecting *ordinary* sorts of beliefs—including one's reasons for moral judgments of others (Haidt 2001), and even mundane reasons for buying one pair of stockings over another (Nisbett & Wilson 1977)! Thus, a person might identify "*p*" as one of her beliefs during armchair self-reflection, even though she does not really believe *p*—and perhaps

⁷ Though it is little noted, the externalism/self-knowledge debate reflects a historically important clash between philosophical paradigms before and after the linguistic (/conceptual) turn. On one hand, Descartes held that one can know one's thoughts independently of knowing the external world. On the other hand, Frege held that content fixes reference (or in Carnap's terms, that intension determines extension). Yet as Putnam (1975, p. 218ff.) argues, Frege's idea seems to lead to externalism: If content *determines* reference, then a difference in reference between Oscar and Twin Oscar shows a difference in content. That holds, even if narrow psychological states are the same. So: If Frege's view implies externalism and precludes Cartesian self-knowledge, then apparently one of these paradigm-defining ideas has to go. Either Descartes was wrong to say we can know contents apriori, or Frege was mistaken that content determines reference. The externalism/self-knowledge debate thus commands our serious attention.

even *denies p*. The conclusions she infers then may deviate from what should be inferred from her beliefs. And thus, reflective reasoning will take her further, rather than nearer, to appreciating what she ought to believe. Call this “Wayward Reflection via Attitude Switching.”

Both versions of the Problem suggest we self-scrutinizers are all in “epistemic bad faith” given what psychosemantics and experimental psychology suggest. Again, critical self-reflection presupposes knowing what one’s beliefs are. So if the mind is largely opaque to itself, why attempt self-reflection? The activity would just show *naivety about psychology*.

Granted, I think there are cases of reflection that are rather sad, due to misjudging one’s own beliefs or desires. I might even be persuaded that in *most* cases, argument amounts to confabulating reasons *post hoc*, to rationalize whatever one has said or done.⁸ And so, what was supposed to be “a way of screening our beliefs in order to make them more accurate turns out, instead...to be little more than self-congratulation” (Kornblith 2012, p. 3). Even so, my hope is to show that self-reflection can be a reasonable activity, and that it is reasonable to believe as much. It is to show that despite findings from psychology, self-reflection can be a rational pursuit.⁹

1.2 Methodological Crisis in Philosophy and Psychology

There is a special case of the Waywardness issue worth singling out, of particular concern to contemporary philosophers. These days, there is much discussion about *intuitions* in philosophical methodology. The usual view is that philosophers rely on intellectual “seemings”

⁸ I have even seen instances of the Dunning-Kruger effect, re: self-knowing. A person claims to be acutely self-aware, yet this ends up being another example of her/his self-ignorance. (A puzzle: How do I know *I* am not such a person?)

⁹ This is not to endorse what Doris (2015) calls “reflectivism,” the view that moral agency requires deliberation on one’s act. I agree with Doris that much less is required. (See also Kornblith 2012.) Yet even if moral agency is not under fire, Wayward Reflection is still worrisome, assuming it deprives us of *some* notable moral/epistemic benefits.

about, e.g., the workings of language, the metaphysics of objects, the ethics of cloning, etc., to arrive at considered philosophical positions. Indeed, the prominent role of intuitions is (arguably) what distinguishes philosophical inquiry from the usual sort of scientific inquiry.¹⁰ Now there is some debate on whether intuitions play the role of *evidence* in philosophy (see Cappelen 2013). But at the least, inquiry has to begin somewhere, and if “intuitions” just concern how things seem at the start of inquiry, intuitions are where we *must* start. (However, this certainly does not mean they have axiomatic status or the like.)

Yet “intuitions” sounds suspicious. Why think one’s intuitions are at all *reliable* in discovering philosophical truths? (For a good overview of the issue, see Nagel 2007.) Wherever you fall on this, however, there is a prior matter—one which is underappreciated in the intuitions literature. Namely, in light of current psychosemantics and experimental psychology, how could a philosopher know what her intuitions *are* to begin with? Nevermind reliability—the more basic question is whether a philosopher can recognize her intuitions in the first place.¹¹ The presumption that the philosopher knows her own intuitions may just show *naivety about psychology*.

¹⁰ There is a case to be made that intuitive judgments are ubiquitous in science also. Williamson (2009) says as much, in discussing Weinberg’s (2009) empirical data against the reliability of intuitions: “Weinberg’s paper is itself full of informal qualitative epistemological judgments, for example about whether we are justified in believing that armchair methods in philosophy are reliable. Nor could any current natural science proceed without such judgments. Even statistical data need to be interpreted; the judgment that they render some hypothesis untenable remains an informal, qualitative one, whatever formal and quantitative considerations it draws on” (p. 474). I doubt the worry about intuitions is comparable for science as for philosophy, yet the issue indeed exists for science to a notable degree.

¹¹ Relatedly, there is a concern about “conceptual analysis:” Nevermind whether conceptual analysis is a reliable method—the prior question is whether one can know about one’s own concepts in the first place.

So Waywardness does not concern merely the type of self-scrutiny that (hopefully) everyone engages in. It also concerns the type of reflection utilized in professional philosophy. But before moving on, let me note a further motive for confronting Wayward Reflection. This should be of particular concern to psychiatrists and professional psychologists, and shall be explored in more detail in chapter 2. Yet briefly, the problem is that experimental method within psychology seems to have the same naivety about the opacity of mind!

Very often, experimental design in psychology gives central importance to subjects' self-reports of mental states. Participants in an experiment are first subjected to some condition, and then are asked to describe something about their own mental state. E.g., a subject is asked what she *thinks* when she sees the ink-blot, what she *believes* of the person shown in the photograph, or what she *judges* to be the best decision in a hypothetical situation. And standardly, the method presumes the subject is able to discern what she herself thinks, believes, or judges in the first place. Yet here is where the Problem arises. In a bit of irony, the evidence gathered *by that very method* suggests subjects are *not* reliable in identifying their own mental states. Thus, there seems to be a *crisis of methodology* in experimental psychology. The method used to gather psychological data rests on an assumption which the self-same data falsify!¹²

Again, this shall be discussed more thoroughly in chapter 2, where we will look in detail at specific psychological studies, to see how this plays out exactly. For now, the point is just that the opacity of mind may present a serious problem for methodology in psychology as well, besides philosophical methodology and self-reflection generally.

¹² By the way, what goes for psychology also goes for experimental philosophy. There too, experimental design gives central evidential weight to self-reports of beliefs, desires, etc. In a further twist, Wayward Reflection also bears on the experimental brand of philosophy that tries to avoid problems with intuitions.

1.3 The Start of an Answer

The book was written to confront these matters. The positive claim will be that, under some recognizable conditions, we can indeed armchair-know our own minds. In fact I will develop the radical view that, in some identifiable circumstances, we are *infallible* about our own occurrent beliefs—judgments about what we judge cannot possibly be false! Even so: I reject the Cartesian idea that higher-order judgments are *indubitable*. (I'm inclined to think the range of what can be doubted is limitless.) Relatedly, I endorse Quine's (1951) view that any belief is subject to empirical revision, if the evidence available warrants it. Even so, if some reflective judgments are *de facto* infallible (and we can recognize when that is so), then that can go a long way toward quelling anxiety about Wayward Reflection.

Before infallibility is dismissed as a Cartesian pipe dream, however, here is a quick and dirty argument for why it has a chance. Consider that we all accept one type of infallibility thesis:

(INF0) Necessarily, if a subject *S* judges that *S* is in mental state *M*, and that judgment is true, then *S* is in *M*.

This of course is trivial. Nonetheless, (INF0) highlights two points. First, it suggests that infallibilism is true if it is restricted to the *right subset* of second-order judgments. Second, and in light of that, it prompts the question whether this subset—or a subset of that subset—can be identified in some other, substantive way. It is not obvious that the answer is “no.” And if the answer is “yes,” then a non-trivial infallibilism holds.

There will be objections of course. But this argument is just meant to soften up the reader. Let me also repeat that the motivation is not to shore up a foundation for empirical knowledge. Again, that sort of Cartesian endeavor seems futile. Rather, infallible self-knowledge is sought in order to counter doubts about reflection—and in particular, about self-reflective critical thinking.

It is meant to vindicate the practice by suggesting that, in some recognizable circumstances, reflection infallibly targets one's own occurrent thoughts and beliefs.

An important caveat: The infallibilism to be defended is *not* seen as an exhaustive view of all self-knowledge. On this, I instead align with Schwitzgebel's (2012) *pluralism*, which holds that most contemporary theories of self-knowledge are right about *some* self-knowing phenomena. (See also Coliva 2016.) Schwitzgebel writes:

[Self-knowing]¹³ is not a single process but a plurality of processes. It's a plurality both *within* and *between* cases: Most individual [self-knowing] judgments arise from a plurality of processes (that's the within-case claim), and the collection of processes issuing in [self-knowing] judgments differs from case to case (that's the between-case claim). (p. 29)

But despite such pluralism, the book shall focus only one type of self-knowledge, viz., the allegedly infallible type. Yet this should not suggest an anti-pluralist stance. I concur that many self-knowing processes can be operative in a given instance (the within-case), and that other self-knowing judgments owe to other processes (the between-case).¹⁴ Nevertheless, I focus on infallibility mechanisms since they yield a particularly striking kind of self-knowledge, one which provides the strongest antidote to Wayward Reflection.

¹³ Schwitzgebel speaks of "introspection" rather than self-knowing as such. But he soon incorporates e.g., "agency-theoretic" and "expressivist" views into his pluralism, even though such views avoid talk of introspection. (See Moran 2001; Bar-On 2004, etc.). (Still, introspection in Schwitzgebel is not the usual sort; it more resembles Evans-style transparency. See especially Schwitzgebel 2011, p.137 along with the attending endnotes.)

¹⁴ Indeed, Part III is rather in line with Carruthers' (2011; 2015) view that self-knowledge is based on interpreting one's own behavior (along with data from proprioception, interoception, etc). Yet while I think Carruthers is right in many ways, I do not think his view respects all self-knowing phenomena. I discuss his view more critically in ch. 2.

1.3.1 Opening objections

There may be a few basic worries about the project as a whole. For one, it can be unclear how any sort of infallibilism can be made empirically respectable.¹⁵ Empirical data shows us only what *is* the case, not what is *necessarily* the case. So how can an evidence-based approach defend the *infallibility* of some judgments? The question is understandable, but fortunately the reply is straightforward. Empirical data can indeed justify *nomological* necessities—including, perhaps, that the occurrence of some judgments nomologically necessitates their truth. Granted, this would not make for the *Cartesian* infallibility of a judgment, defined as truth in all epistemic/doxastic possible worlds (or the like). But as we shall see, an empirically-minded philosopher can support that some judgments are true whenever executed, as a matter of psychological law. This would make them “infallible” in one sense, albeit not in the Cartesian sense.

A different objection is that the project is undermotivated. Some may feel that reflection holds enough intrigue, whether or not it targets one’s own beliefs. E.g., a philosopher may be undecided on modal fictionalism, but she can be motivated to engage the arguments regardless. This of course is quite right. Yet for one, she would likely assume some armchair self-knowledge, viz., that she regards fictionalism as a live option at least. But more importantly, if reflective inquiry is to have anything more than entertainment value, it must be *rationally* motivated, and not just accidentally amusing. In particular, it must be reasonable to believe that reflection can achieve its goals, including that of self-evaluation. So we need reason to think that, at least in some recognizable circumstances, the beliefs reflected upon are indeed one’s own.

The endeavor of the book also faces a kind of self-referential objection. My plan is to use armchair reflection to defend armchair reflection. But this seems viciously circular. This can be

¹⁵ My thanks to an anonymous referee for raising this concern.

elaborated in two ways (corresponding to the two versions of the Waywardness Problem). First, if content externalism is true, then my reasoning may equivocate in ways that I cannot detect from within the armchair. Thus, premises about reflection may be used to infer conclusions about “twin-reflection” without my realizing it. This may strike you as idle skepticism, however. Yet I do not wish to be so dismissive toward the objector. But the matter cannot be settled quickly, and it is best addressed in chapter 6, where I discuss externalism and equivocal reasoning more broadly (see esp. the final footnote of the chapter).

The second version of the self-referential charge is that, per experimental psychology, I cannot know whether my stated assumptions on critical reflection capture my own judgments.¹⁶ But to this, I say “fair enough.” I will proceed in ignorance of whether various claims express what I believe.¹⁷ I shall even feign neutrality on whether a view strikes me as a live possibility. (I will just be agnostic on why I find some proposition worth thinking about.) Fortuitously, I am still interested enough in the issues that I don’t need a *rationalizing* motive to get started.

Though again, if reflective inquiry is to have anything more than entertainment value, the hope is that reflective inquiry can be rationally vindicated. And in attributing this hope to myself, I am perhaps begging a question. But my reasons for *interest* in the issues can be separated from my arguments on the first-order issues themselves. So there need not be any question begging at the ground level. (I may inevitably beg a question concerning why I should care about the arguments. Yet that is not to say that the arguments themselves beg.)

¹⁶ Such circularity, I take it, is closely related to Kornblith’s (2012, ch. 1) regress problem.

¹⁷ I may occasionally speak of “what I believe” for rhetorical or stylistic reasons. But the point is that no argument will depend on any premise about what I, TP, judge to be one of my thoughts/judgments.

*1.3.2 *Internal- and external-world skepticisms*

But admittedly, there remains a more difficult aspect of the self-referential objection. Consider: If my *modus operandi* is largely philosophical, then per section 1.2, I apparently must draw upon philosophical intuitions in reasoning about various claims. And that requires recognizing what my intuitions are. Yet the ability to recognize one's own intuitions is partly what's in doubt. So if the project of the book requires such an ability, then there indeed seems to be a worrisome circularity built into the endeavor.

Of necessity, the reply here is more programmatic—for it is meant to counter a more programmatic kind of skepticism. Concurrently, as an expository tactic, it proves useful to consider briefly the problem of external-world skepticism.

In an ordinary debate, a question about question-begging can be perfectly apt; for ordinarily, the arguments on each side should start from common ground. But things take an odd shape when debating the external-world skeptic. One such skeptic, by definition, concedes no ground about the external world at all. The shared ground is thus invariably insufficient for vanquishing him. So, simply by being sufficiently uncooperative, the skeptic remains undefeated.

Mooreans thus ignore this skeptic, insisting that anyone *that* skeptical is intellectually disingenuous, if not insane. (And one *should* ignore the disingenuous or insane!) Regardless, our skeptic can teach us something. The lesson is apparently that one cannot have reasons for beliefs about the external world, without making presumptions about the external world. Let us suppose, nonetheless, that this does not threaten our knowledge of the external world. Still, the skeptical lesson at least raises an explanatory challenge for, e.g., cognitive science. Namely: *Given* that we know the external world, partly on the basis of mere presumptions, *how* do we manage to do that? By what means are we able to know in this presumptuous manner?

In focusing on this explanatory question over traditional skepticism, I show an allegiance to Quine's (1969) *naturalized epistemology*. Yet in the present context, Quine's program is understood a bit differently than is typical. After all, the explanatory question here is not answered by studying the neurology of the brain or the various perceptual mechanisms (though that may be relevant). For what animates the question is a puzzle about our *reasons* for belief, and reasons do not show up on an fMRI scanner (at least, not *qua* reasons). The puzzle again is how we can know about the world despite our reasons being oddly ungrounded or circular. To repeat, it is taken for granted that we *do* know; the question is merely *how*, given the structure of our reasons.¹⁸

For my part, Sellars (1956/1963) answered this in a single, brilliant brushstroke: “[E]mpirical knowledge, like its sophisticated extension, science, is rational, not because it has a *foundation* but because it is a self-correcting enterprise which can put *any* claim in jeopardy, though not *all* at once” (p. 170, his italics). Here, the “foundation” is basically the common ground that we would share with the skeptic. It would consist in certainties, beliefs that one is unable to doubt regarding the external world. The explanatory puzzle arises since we tend to assume that, to know the world, we need non-circular reasons starting from such common ground. But by definition, our skeptic does not concede enough ground for that. Thus, the puzzle.

Sellars teaches us, however, that this foundationalist conception of knowledge is a mistake. We can suppose that knowledge is true rational belief, yet rational belief need not be *skeptic-immune* belief. It is instead enough if the belief is situated within a “self-correcting enterprise,” a system of beliefs where nothing is adopted dogmatically. Self-correction means

¹⁸ I take this to be a puzzle only about *reflective* knowledge, not about so-called *animal* knowledge. Cognitive neuroscience alone might very well explain animal knowledge. (The distinction is from Sosa 1985, 2011, etc., yet I would revise Sosa's way of drawing the distinction a bit, in light of sound criticism from Kornblith 2004, 2016.)

“checking” one belief against other beliefs, presumably by how well the one coheres with the others. Accordingly, if a belief creates enough conflict, it is removed. The proposal, then, is that such coherentism is the correct account of what makes a belief “rational,” and shows how true rational belief possible.¹⁹

Sellars’ idea still leaves some anxiety. Most notably, it leaves us with (what Bonjour 1985 calls) the “alternative coherent systems” objection. The objector notices that, in principle, vastly different systems could each count as having fully justified or rational beliefs as members. If all that’s required is a certain kind of coherence among the members, that hardly picks out a unique system. This, consequently, raises the concern that a belief’s being “rational” is ultimately not a *truth-conducive* property—it is unclear whether that feature increases the objective likelihood of a belief being true.²⁰ However, the intuition is that justification must render a belief more objectively probable. After all, we want “justification” to corroborate our choice of belief, vis-à-vis the goal of attaining truths and avoiding falsities.

¹⁹ On coherentism, see Quine & Ullian (1970/1978), Sellars (1973), Rescher (1973; 1979), Lehrer (1974; 1989a, b; 2000; 2003), Bonjour (1985; 1989), Harman (1986), and Lycan (1988; 1996; 2002; 2012). See also Foley (1987), who explicitly speaks of rational rather than justified belief. I first became worried about Wayward Reflection in thinking about Bonjour’s (1985, p. 101) “Doxastic Presumption.” Yet strictly speaking, the Presumption concerns an individual’s grasp of the *entire* belief-system, and such global coherentism seems unattractive (see n. 21). But localized coherentism would still need to contend with Waywardness.

By the way, “checking” a belief against other beliefs admittedly engages in the circularity as well, re: self-knowing. For the process requires knowing what these other beliefs are. (The circle also exists in knowing your intuitions about what coheres with what.) Ultimately, the reply to internal-world skepticism is meant to address all instances of the circle.

²⁰ The most powerful critique of this sort is from Klein & Warfield (1994).

Without going into detail, the objection is essentially stating the underdetermination of theory by evidence, as made famous by Duhem and Quine. Quine (1951) puts the matter thus:

The totality of our so-called knowledge or beliefs...is a [hu]man-made fabric which impinges on experience only along the edges... A conflict with experience at the periphery occasions readjustments in the interior of the field... But the total field is so undetermined by its boundary conditions, experience, that there is much latitude of choice as to what statements to re-evaluate... No particular experiences are linked with any particular statements in the interior of the field, except indirectly through considerations of equilibrium affecting the field as a whole...Any statement can be held true come what may, if we make drastic enough adjustments elsewhere in the system...Conversely, by the same token, no statement is immune to revision. (pp. 39-40)

It may overstate matters to say that the *whole* system is implicated when encountering recalcitrant experience (cf. Quine 1991).²¹ However, it is widely accepted that our theories of the world are undetermined by the evidence, along the lines Quine indicates.

So it seems the “alternative coherent systems” objection is not to be resisted as much as admitted, as an important philosophical truth about our epistemic position. Yet the objection stands that a belief’s being “rational” now fails as a truth-conducive property. In reply, however, I follow Quine who construes the rationality of belief not by its connection to *truth*, but rather in *pragmatic* terms:

²¹ Relatedly, at no point do I say that to properly *acquire* beliefs, the epistemic agent must be guided (consciously or not) by considerations of global coherence. I concur with Kornblith (1989) that such things require a computational complexity that is unrealistic. Even a coherentist “check” of a belief can operate more locally, using only a proper subset of domain-relevant beliefs. (This is what Kornblith 1989 himself accepts in his Quine-style view of confirmation; see p. 212.)

Each man [sic] is given a scientific heritage plus a continuing barrage of sensory stimulation; and the considerations which guide him in warping his scientific heritage to fit his continuing sensory promptings are, where rational, pragmatic. (1951, p. 43)

Pragmatic desiderata include maximizing “super-empirical virtues” like simplicity, conservativeness, fertility, scope, etc. Such concerns can guide a rational decision among “alternative coherent theories,” even though the super-empirical virtues might not always be truth-conducive. Yet if they allow for rational theory-choice regardless, knowledge again seems possible, despite the circular structure of our reasons. And the point would remain that knowledge does not require a skeptic-immune foundation, but rather just a place in a system that implements coherentist self-correction, where the rationality of the whole system ultimately has to do with its pragmatic advantages.

What then of truth and falsity? “Justification” that is not truth-conducive can sound like a contradiction in terms. But this may just indicate we should follow Sellars and Quine in speaking of “rational” rather than “justified” belief. Still, if the rationality of a belief is largely orthogonal to truth, then our means to maximizing truth and minimizing falsity seems not very effective. (Prisoners of The Matrix can just as easily achieve fully “rational” belief, despite being systematically mistaken about the outside world.) However, rational belief is not wholly independent of truth, since consistency is the minimum required on a true theory. (I would argue that that ontological parsimony, judiciously applied, helps us avoid falsity as well. Perhaps the success of science also assures us that we are not entirely off track.) Regardless, a less-than-effective means to truth may be our lot in life, given the ungrounded or circular structure of our

reasons. And perhaps truth is not the goal of inquiry as much as a regulative ideal. The more immediate goal, apparently, is to maximize coherence and the super-empirical virtues.²²

The preceding sketch is just a sketch; even so, it contains much that is controversial. However, I cannot pursue the debates at this time. The Quine-Sellars view at least offers a minimally tenable way for responding to external-world skepticism (although it is admittedly fairly concessive). Nevertheless, if knowledge is a species of true rational belief, it still reveals how knowledge of the external world is possible, so that despair is not inevitable.

My suggestion, furthermore, is that the view also indicates a minimally viable tactic against skepticism about the *internal* world. Our internal-world skeptic is similarly pointing out a circularity in our reasons—that one must presume access to one’s intuitions to show how it is possible to access (inter alia) one’s intuitions. But if we extend the Quine-Sellars view, this circularity need not preclude that some beliefs remain more rational than others, insofar as some beliefs better cohere with the others (in a self-checking system). For coherence makes for rational belief at least in a pragmatic sense of ‘rational’, and also in a sense not wholly orthogonal to truth.

Just to be clear, coherentism was not invoked to explain reasons for belief in the “context of discovery.” Typically, when I recognize that I am hungry, my reasons do not relate to how well “I am hungry” coheres with my belief-system. The discovery is usually instead based on introspective evidence (yet per Schwitzgebel’s pluralism, it can be based on other things as well). Instead, I appealed to coherentism only in a specific “context of justification,” namely, that of

²² This may lead some to conclude that maximizing truth and minimizing falsity cannot *really* be our goal in inquiry. The matter likely depends on what “the goal of inquiry” means exactly. But there is a sense, I think, in which inquiry typically has truth as a goal (perhaps inter alia). Though whether that *should* be a goal is yet a further question which I do not address here.

combating internal-world skepticism.²³ (Though coherence likely plays a role in other contexts as well.)

We must acknowledge an important disanalogy between internal- and the external-world skepticism. Namely, the internal-world skeptic has *non-idle* reasons for doubt, viz., reasons gleaned from psychology. Thus, compared to the external-world skeptic, the internal-world skeptic has a *better* case for suspending judgment about her domain of interest. This is why, as in chapter 2, I call her not a “skeptic,” but rather an “empirically-based antagonist” toward self-knowledge or knowledge of the internal world. (Though n.b., in chapter 2, antagonism concerns only a particular kind of self-knowledge, viz., the supposedly infallible kind. But see chapters 10 and 11 on antagonism more generally.) Thus, even if the Quine-Sellars view shows how one can know the external world, does it adequately buttress the possibility of *self*-knowledge?

There is a mistaken presupposition in the question. Strictly speaking, it is not simply coherentism-*cum*-pragmatism that shows how we might have external-world knowledge. If we are being careful, that view just *sets the terms* on what would count as such knowledge. But it does not yet follow that our beliefs *satisfy* those terms. An analogy: ‘*x* is round & *x* is square’ defines what it is to be a “round square,” but it does not follow that some object satisfies that definition. In the case of the external-world knowledge, of course, it is natural enough to say that our beliefs often *do* satisfy those conditions. In contrast, however, beliefs about one’s internal world may not cohere adequately with other beliefs, specifically, beliefs about psychological findings.

²³ ‘Justification’ as it appears in ‘the context of justification’ is used to denote a *process* (“the game of giving and asking for reasons”). This, by the way, affords a clear distinction between the two types of context: When acquiring new evidence, one can discover *p* (= know *p* for the first time) *sans* a process of justifying *p* to someone else. Yet *continuing* to know *p* may depend on your ability to quell doubt in subsequent contexts.

So in brief, when antagonism's "evidence-based" feature is stressed, it is unclear if beliefs about the internal world qualify as knowledge. Nevertheless, the psychological results are not a direct challenge to the conditions themselves. (For that, normally one offers a counterexample or the like.) The Quine-Sellars idea *per se* is not affected by stressing the empirical difficulties.

But no doubt, the results from psychology feed antagonism toward self-knowledge. Again, knowing these results, our beliefs about the internal world seem not to cohere sufficiently with the rest of the system. The coherentist is thus led to ask about the rationality of these beliefs. The eventual goal, however, is to show that hypothesizing self-knowledge is not just permissible, but also can be made rationally respectable, despite what psychology tells us. So although I cannot cut off antagonism at this point, that will be the ultimate aim in what follows.

1.3.3 Epistemic internalism?

There is a follow-up objection, however, and it is a broader concern for the book as a whole. The sense may be that *epistemic internalism* is being assumed, a view where knowledge requires evidential reasons to be introspectively accessible. After all, the previous sub-section favors a kind of coherentism (which is usually typed as internalist)—and besides, the book's central concern is with armchair self-knowledge. However, epistemic internalism is highly controversial. Apparently, it precludes reliabilist epistemology (cf. Goldman 1986), along with various kinds of social epistemology currently on the scene (cf. Goldman 1999; see also the papers in Goldman & Whitcomb 2011).

Yet I myself am a social epistemologist. To me, it is obvious that even Descartes needed peer-feedback on *The Meditations* before he could see it as establishing anything. (The isolated individual is just too unreliable, especially in philosophy.) Thus, I hardly mean to presuppose

epistemic internalism.²⁴ But the subsequent chapters do not assume epistemic externalism either. Neutrality on the matter is possible since, surprisingly, the book is largely unconcerned with the evidential reasons that enable self-knowledge

In this respect, ‘self-knowledge’ may be misleading. After all, the book’s basic contention is that specific kinds of second-order judgment are *infallible*. That is to say, these judgments have a distinctive *semantic* feature—they are invariably *true* whenever executed. Given that this thesis is semantic, one could see the book as more a study in the philosophy of mind than in epistemology. (That is so, especially since the ancillary themes are content externalism, methodology in psychology, and the possibility of eliminativism.)

Nonetheless, ‘self-knowledge’ is not truly a misnomer, for chapters 5 and 10 argue that infallibility can help secure *knowing what* occurrent mental state I am in. Yet knowing-what is markedly different from the traditional topic of epistemology, knowing-that. For whether true belief suffices for “knowing what” depends crucially on one’s contextually-determined *goals* or *purposes*—specifically, on whether one’s beliefs are adequately *informative* for those purposes. To be sure, knowing-what also depends on one’s reasons. And in these chapters, I explain how it is possible to armchair-justify second-order beliefs to either a Cartesian skeptic (chapter 5) or to

²⁴ Still, I see the appeal of internalism. Per the Preamble, the individual faces “alone” the question “What should I believe?,” and internalism seems apt in that setting. Given such epistemic responsibility, I thus balk at Kornblith’s (2013) verdict that the internal perspective “is *wholly* underserving of the special place which...philosophers would give it,” (p. 126, my italics). Yet if the individual is responsible for answering the question, her *methodology* of answering need not be individualistic. She is usually better off conferring with others. (I thus also *agree* with Kornblith 2012, chs. 2 and 3, that the role of individualistic inquiry within epistemology has been overblown.)

an evidence-based antagonist (chapter 10).²⁵ However, at no point do I *require* introspective access to one's reasons in order to know-what. But nor is such a requirement dismissed. Neutrality on the internalism question is therefore sustained.²⁶

As for the coherentism invoked, I can similarly feign neutrality on whether it is an internalist or externalist view. Note well, it is quite possible to be an externalist coherentist. At least, the "beliefs" in question might be those of a group, rather than of the individual.²⁷ A social coherentist might even invoke the reliable "cognitive processes" used by the group in arriving at its beliefs, e.g., peer review, the replication of experiments, etc. (Indeed, this probably how to best understand the Sellarsian "self-correcting enterprise")

Finally, the book's interest in armchair self-knowledge does not require assuming epistemic internalism. For one, a process-reliabilist can take an interest, given that some reflective

²⁵ I think of justification here as a social activity (vs. a static property), though I will not demand this in what follows. One may object that a social epistemologist should deny the possibility of justification if Cartesian skepticism is live. For then, justification should be possible even if no one else exists. But in fact, "justifying" remains a social activity even if it involves only two individuals (a skeptic and myself). Yet I would resist the idea that in an *utterly solipsistic* world, I can "justify" beliefs to my own imaginary skeptic. (I'm not even sure what that would mean.)

²⁶ In chapter 5, the reasons attending an infallible self-attribution take the form of introspective judgments—but again, I do not *require* a role for introspection in self-knowing. In chapter 10, however, the relevant reasons end up being *perceptual* rather than introspective. Yet this too is compatible with either epistemic internalism or externalism. Internalism says not that one's reasons must be *introspectings*; rather, it must be possible for one's reasons to be *introspected*. And presumably, both perceivings and introspectings can be introspected.

²⁷ See Huebner (2013) for a stellar defense of group cognition. Social coherentism can seem dubious given that coherentism usually requires introspective access to reasons. And "group introspection" is obscure at best. But again, coherentism here does not require introspectively accessible reasons for rational belief. Though granted, this means it is not the most familiar sort of coherentism.

judgments are defended as highly reliable (*nay, infallible*). Second, it will emerge in Part III that linguistic communication enables a crucial type of self-knowledge—and communication in the knowing-process is of central interest to social epistemology. Third, a key concern of the book is to show that externalism about content is compatible with armchair self-knowledge. This bears important connections with social epistemology, for it concerns the most basic ingredient of knowledge—thought—and how that ingredient is shaped by the community. Indeed, if one is an externalist about thought-content, then one is an externalist about knowledge-content. And so to that extent, a content externalist *is* a social epistemologist.

1.4 Overview of the Book

Often an introductory chapter contains what are, in effect, abstracts for the chapters. But here, the abstracts have been written as abstracts, and can be found before the preface. In what remains of this chapter, the aim is instead to describe the book from a bird's eye point of view, to give a sense of what (if all goes well) the chapters achieve collectively.

The book is divided into four parts (although Part IV is just a summarizing and concluding chapter). Part I consists of preparatory work for what happens later; it includes the preamble, this chapter, and the next. Part II is designed to counter Wayward Reflection via Thought Switching, while Part III is meant to allay Wayward Reflection via Attitude Switching. The layout of Parts II and III are parallel: Each begins with one or two chapters articulating some infallibilist thesis about second-order judgment, followed by two chapters addressing objections, and ending on a chapter about how the newly minted infallibilisms subvert Waywardness (chapters 6 and 11, respectively).

At bottom, the issue is whether a person can know what she judges (= occurrently believes), for typical self-reflecting, self-reporting (e.g., in a psychological experiment), or philosophical

inquiring. Throughout, it is assumed that in order to self-know what one judges, two things must be in place. First, it must be known what *thought* is being judged, or to say much the same, what the *content* of one's judgment is. Second, it must be known that one has the *judging-attitude* toward that thought (or, as I like to call it, the "alethic pro-attitude")—as opposed to the wishing-attitude, the doubting-attitude, the denying-attitude, etc.

With this in mind, Part II defends infallibilism about self-knowledge of thought, whereas Part III details an infallibilism about self-knowledge of judgment. For those in the know, both parts take after Tyler Burge's (1988) view, where knowledge of one's occurrent thoughts is "self-verifying". (See also Davidson 1987 and Heil 1988.) However, Burge's account can seem open to serious objections, and he does not extend the "self-verifying" idea to self-knowledge of attitude. These matters are remedied in Parts II and III, respectively. Such infallibilist views are then shown to help stall (the respective versions of) the Waywardness Problem. The two Parts thus jointly support the possibility of armchair-knowing one's own judgments, and thus of rationally self-reflecting, despite contraindications from psychology.²⁸

It is worth remarking that, besides this chapter, chapter 2 is a "must read," for it shows that some form of infallibilism about self-knowledge is consistent with the existing psychological data. This is crucial, since many will antecedently think that the empirical evidence debunks such a view. Not so, as we shall see. Yet one may want to read the preamble chapter prior to that. It concerns doubts about *philosophy* in general. But beyond this, the preamble could be skipped, as its details

²⁸ It may be said that self-reflection also requires knowing one's own desires, intentions, etc. Then, even if the book upholds self-knowledge of judgment, the rationality of reflection still has not been shown viable. But for one, the accounts of chapters 7 and 8 can be extended to self-knowledge of other mental states (and I mention how in the appendices of each chapter). Second, paradigm reflection just consists in inferring one judgment from others. So if self-knowledge of judgment is shown possible, that at least establishes that paradigm self-reflection is possible.

shall not be crucial elsewhere. Nonetheless, the preamble is essential to grasping the larger, metaphilosophical orientation of the book (which I myself see as important).

As the final matter, I might stress the importance of chapter 8 against Wayward Reflection. For a select class of judgments, the chapter defends an infallibilist variant of a “Neo-Expressivist” view (cf. Bar-On & Long 2001, Finkelstein 2003, Bar-On 2004). Yet it borrows from other views on the scene as well, including constitutivism (Shoemaker 1996, Wright 1998, Coliva 2009), “agency-theoretic” views (Bilgrami 2006, O’Brien 2007, Soteriou 2013), and “transparency” views (Byrne 2005, Fernández 2013).²⁹ However, this is where the book is most engaged in the philosophy of language, which generally tends to be more difficult. But chapter 8 is at the heart of the book, and my belief is that the attentive reader will be rewarded.³⁰

²⁹ In such ecumenism, the chapter 8 account has the closest affinities to Moran (2001) and Heal (2002). Though unlike myself, none of these authors defend any sort of infallibilism.

³⁰ I suspect that Part III is in many ways the most impressive part; it is unfortunate that exposition forces it to occur third. But its basics can be grasped *sans* Part II, and for the time-pressed reader, I might advise visiting it first.