ON KNOWING YOURSELF AND BEING WORTH KNOWING

Jordan Louise MacKenzie

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements of the degree of Doctor of Philosophy in the department of Philosophy.

Chapel Hill
2017

Approved by:
Thomas E. Hill
Susan Wolf
Simon Blackburn
Ram Neta
Matthew Kotzen
ABSTRACT

Jordan Louise MacKenzie: On Knowing Yourself and Being Worth Knowing
(Under the direction of Thomas E. Hill)

Why should you know yourself?

While philosophers have paid ample attention to the epistemological issues surrounding the Delphic Command to ‘Know Thyself’, they have been comparatively quiet on the question of why we ought to care about heeding it. When attention has been paid to the question of why we ought to know ourselves, the answers that it has generated have typically been broadly instrumentalist. Self-knowledge is instrumentally valuable in many ways: we are better moral agents for not being self-deceived, and we are better rational deliberators for knowing what we desire.

These sorts of instrumentalist explanations, while compelling, do not fully explain the value that many of us place in self-knowledge. We demand the truth, even when the truth hurts, and we try to understand our motivations for long-ago actions that have few, if any, practical implications for our present-day selves. This non-instrumental commitment to self-knowledge is not just common, but also admirable. There is, after all, something impressive about people who want to see themselves ‘warts and all’.

In my dissertation, I explain why we have reasons to value self-knowledge that are independent of the valuable consequences that self-knowledge often helps bring about. I argue that, just as interpersonal love gives us non-instrumental reasons to know its objects, so too does self-love give us non-instrumental reasons to know ourselves. Further, while
interpersonal love is not something that we can owe to other people, self-love is something that we can owe to ourselves. This is because self-respect requires us to act in ways that adequately embody our values and honor our commitments. In doing so, we extend to ourselves the sort of partiality that is required by self-love, while at the same time making ourselves into people whom we can love. Thus, to see oneself as worth knowing for the sake of knowing is part of what it is to love oneself, which is something that self-respect demands that we strive to do. To pursue self-knowledge for its own sake, in turn, is to respectfully honor one of the commitments that self-love motivates us to make.
ACKNOWLEDGMENTS

Tom Hill has been a patient, thoughtful, and supportive supervisor who has helped me to develop my own philosophical voice. I am so lucky to have had his supervision on both this project, and on my Master’s thesis. I would also like to thank Susan Wolf, Simon Blackburn, Ram Neta, and Matt Kotzen for their incisive criticisms, helpful comments, and unflagging support.

During the past six years, I have also benefitted from countless conversations with fellow graduate students and faculty members. In particular, I would like to thank Krasi Filcheva, who has been an incredible philosophical interlocutor throughout the entire dissertation-writing process. Thanks also to Sam Reis-Dennis, Lindsay Brainard, Jay Brantner, Benjamin Bagley, Dana Falkenberg, John Phillips, Chris and Joanna Blake-Turner, Keshav Singh, Samantha Wakil, Samantha and Isaac Arten, Doug MacLean, Geoff Sayre-McCord, Michael Cholbi, Kerah Gordon-Solmon, and Rahul Kumar. I am also grateful for the financial support that I received from the Social Sciences and Humanities Research Council of Canada.

Most of all, I must thank my family. To my mother, sister, and nanny, thank you for supporting my decision to take a gamble on philosophy. To my dog, Rosie, thank you for forgoing numerous belly-rubs and walks in the service of this dissertation. To my husband, Ted, thank you for reading countless drafts, serving as a constant sounding board, and never
once questioning my ability to finish this project. And finally, to my father, Peter MacKenzie, thank you for introducing the six-year old me to a thing called ‘philosophy’: I wish you were still here to read this.
# TABLE OF CONTENTS

**INTRODUCTION: WHAT FLORENCE DIDN’T KNOW** ................................................................. 1

Section 1: What Type of Self-Knowledge Is (Possibly) Worth Having for its Own Sake? . 6

Cassam’s Characterization of Substantial Self-Knowledge ...................................................... 7

An Alternative: Substantial Self-Knowledge and Identification ............................................. 13

Section 2: The Instrumental Value of Substantial Self-Knowledge ........................................ 19

The Moral Self-Improvement Account ..................................................................................... 20

The Authenticity Account ......................................................................................................... 21

The Well-Being Account ......................................................................................................... 24

Section 3: Against the Explanatory Exhaustiveness of the Instrumentalist View ............ 25

The Instrumentalist View: Conclusion .................................................................................... 28

Section 4: The Ineffability of Intrinsic Value ......................................................................... 30

Section 5: The Structure of This Dissertation ....................................................................... 35

**CHAPTER 1: SELF-KNOWLEDGE AND EPISTEMIC SELF-RESPECT** ............................... 38

Introduction ............................................................................................................................. 38
Section 1: Respect ............................................................................................................................................. 39

Section 2: Respect, Practical and Epistemic ...................................................................................................... 45

Epistemic Rationality and Moral Demandingness ............................................................................................... 52

Section 3: The Requirements of Epistemic Respect ............................................................................................. 56

Section 4: The Requirements of Epistemic Self-Respect ..................................................................................... 62

Two Philosophers ........................................................................................................................................... 67

Felix ................................................................................................................................................................. 69

Sarah ............................................................................................................................................................... 72

Conclusion: Epistemic Self-Respect and Self-Knowledge .................................................................................. 73

CHAPTER 2: SELF-KNOWLEDGE AND THE REASONS OF LOVE .............................................................. 77

Introduction ....................................................................................................................................................... 77

Section 1: Clarificatory Remarks ..................................................................................................................... 78

Section 2: Love’s Reasons ............................................................................................................................... 79

Section 3: Self-Love’s Reasons ...................................................................................................................... 87

What It Means to Relate to Ourselves .............................................................................................................. 88

Why Our Relationship with Ourselves Demands Self-Love .......................................................................... 90

Section 4: From Self-Love to Self-Respect ..................................................................................................... 97
Conclusion: The ‘Substance’ of Substantial Self-Knowledge ........................................... 101

CHAPTER 3: SELF-DECEPTION AS A PERVERSION OF VALUING ......................... 105

Introduction ........................................................................................................................... 105

Section 1: The Desiderata of a Successful Account.......................................................... 107

Desideratum 1: Moral Relevance ....................................................................................... 107

Desideratum 2: Moral Uniqueness ..................................................................................... 107

Desideratum 3: Moral Content ......................................................................................... 109

Desideratum 4: Range of Reactive Attitudes .................................................................... 110

Desideratum 5: Exceptions ............................................................................................... 110

Desideratum 6: Coherence with Epistemological Theories of Self-Deception............. 111

Section 2: What’s So Wrong With Self-Deception?......................................................... 112

The Autonomy Account ..................................................................................................... 112

The Moral Harm Account ................................................................................................ 115

Section 3: Self-Deception, Self-Respect, Self-Love ......................................................... 120

Self-Deception as Epistemic Self-Disrespect ................................................................. 120

The Motivationalist and the Intentionalist ..................................................................... 121

The Epistemic Failings of Self-Deception ....................................................................... 126
Section 4: Self-Deception and Self-Knowledge................................................................. 131

Section 5: Self-Deception as a Perversion of Valuing....................................................... 136

Taking Our Ends Seriously.............................................................................................. 136

Section 6: Meeting the Desiderata of a Successful Account........................................... 142

CONCLUSION: WHAT FLORENCE SHOULD HAVE KNOWN ........................................ 148

BIBLIOGRAPHY................................................................................................................. 151
INTRODUCTION: WHAT FLORENCE DIDN’T KNOW

Florence Foster Jenkins was a New York City socialite renowned for two things: her unflappable love of music and her complete lack of musicality. This lack of musicality was apparent to everyone, with the unfortunate exception of Florence. Indeed, Florence fashioned herself to be quite a songbird, and throughout the 1910’s, 20’s, 30’s and 40’s, was known to ‘treat’ select audiences of friends and admirers to her tuneless warbling in private recitals.¹ These concerts were organized by her long-time romantic partner and business associate, St. Clair Bayfield, who took ample precautions to ensure that his dear Florence never found out that people only came to these recitals for a good laugh.² Florence’s career, if you could call it that, culminated in a solo recital at Carnegie Hall, where her musical styling caused uproarious laughter and general pandemonium. The sound of her voice can be summed up in one pithy line from Earl Wilson’s New York Post review of her Carnegie Hall debut: “Mrs. Florence Foster Jenkins…can sing anything but notes.”

This review gave Florence a rare opportunity to inquire into the truth about her musical ‘talents’. Unlike at her recitals, St. Clair hadn’t been able to keep the music


reviewers out of Carnegie hall, nor had he been able to bribe them into silence.\(^3\) Now, as a matter of fact, Florence did find out about the bad reviews, the historical record is unclear about whether Florence was devastated, or simply dismissive about the reviewer’s comments.\(^4\) But I want to pause here at the moment before Florence found out, and ask: what reason, if any, would she have had she have to want to discover this new information about herself? And what, if anything, would she have stood to gain by learning that she’s a bad singer?

This dissertation, in essence, seeks to provide answer to these two questions. In providing this answer, I will argue that we ought to be committed to knowing ourselves not simply because self-knowledge can be instrumentally valuable, but also because it can be worth having for its own sake.

Before I do this, however, I will first have to complete some philosophical ground-clearing. While I think that Florence would have gained something from discovering the truth about her musical ‘talents’, this intuition has not been shared by many of the philosophers who have written on the value of self-knowledge. When philosophers talk about the value of

---

\(^3\) There are a few exceptions to this. Over the years, reviews of several of Florence’s recitals had appeared in print. These reviews were, on the whole, quite positive, albeit often tongue-in-cheek. A review of her 1935 recital at the Ritz-Carlton, for instance, described a “wildly applauding audience” that “left no doubt of the enjoyment derived by the throng.” Bullock, *Jenkins!!!*, 90–91.

\(^4\) St. Clair Bayfield reported that Florence was “crushed” by the reviews. Bullock, *Jenkins!!!*, 119–20. Francis Robinson, a friend of Florence who later wrote the liner notes to the first CD recording of Florence’s music, disputes this, and claims instead that Florence died with a “happy heart.” Francis Robinson. Liner notes to *The Glory (????) of the Human Voice*. RCA LM 2597-C. Further, as Bullock (120) notes, Florence saved the Carnegie Hall reviews in the same scrapbook in which she saved all of her other reviews. The nature of Florence’s reaction may have some bearing on the circumstances of her death. Florence suffered a heart attack five days after her recital, and ultimately died a month after that. It is unclear whether the heart attack was brought on by emotional distress, or whether her death can be attributed to the fact that she was a 76 year old women who had been in poor health for decades.
self-knowledge, they typically do so in explicitly instrumental terms. Self-knowledge can be a means to moral virtue,\(^5\) autonomous self-governance,\(^6\) and well-being,\(^7\) while its absence opens the door to viciousness,\(^8\) and frustrates our ability to get what we want.\(^9\) If self-knowledge is only valuable for the consequences it helps us secure, I think we can intuitively conclude that Florence had no reason to learn about her lack of musicality. Given her age and circumstances, this knowledge couldn’t have led her to improve upon herself, or to choose a more suitable life’s project. Nor was it likely to positively impact her well-being: indeed, we can reasonably assume that this knowledge would have had the exact opposite effect.

---


There is much to recommend the Instrumentalist View. For starters, it seems plainly correct: we do need self-knowledge to achieve a host of independently-valuable ends. And the view can also be motivated from the opposite direction, by comparing it favorably to the sort of non-instrumental picture of self-knowledge’s value that I will ultimately defend. For some, the idea of wanting self-knowledge for self-knowledge’s sake calls to mind the image of Narcissus endlessly gazing at his reflection in a pool of water. Surely, the value that Narcissus put in knowing himself was no sign of virtue. Rather it was simply, as his name implies, a symptom of his self-obsession. And so we should worry that we only value knowing ourselves for the sake of knowing ourselves because we, like Narcissus, are prone to arrogance and self-obsession. Perhaps, then, we should to pay less attention to ourselves and more attention to everyone else.

We might also wonder whether ascribing a non-instrumental value to self-knowledge commits us to the implausible position that all self-knowledge has this sort of valuable. Presumably, this isn’t the case: it is one thing to think that knowledge about your talents, values, and commitments is intrinsically worth having, but it is another thing to think that knowledge of the number of hairs on your head, or the number of bowls of cereal you consumed last year, is worth having for its own sake. While the Instrumentalist View has a clear way to delineate valuable self-knowledge from valueless self-knowledge, a view that simply ascribes a non-instrumental value to knowing oneself will have a much harder time explaining why not all self-knowledge is worth having for its own sake.

A non-instrumental account of self-knowledge’s value will have to explain why ought to be particularly concerned with our self-knowledge, rather than with somebody else’s. If Florence stands to gain something non-instrumentally valuable from the discovery of her
unmusicality, would I also stand to gain something non-instrumentally valuable by discovering this fact about her? Once again, the Instrumentalist View has a ready response to this challenge: self-knowledge often matters to me in a way that interpersonal knowledge does not precisely because it, unlike most interpersonal knowledge, is instrumental in helping me achieve certain independently valuable ends. A non-instrumentalist account of self-knowledge’s value will have to explain why our self-knowledge matters to us in a way that it does not (and often, should not) matter to others.

My project in this introductory chapter is threefold. First, I aim to get clear on what type of self-knowledge is worth having for its own sake. I’ve already said that knowledge about one’s talents, values, and commitments falls into this category, and that knowledge about one’s cereal consumption does not, but I aim to provide a rough set of criteria that can be used to distinguish what Quassim Cassam has called ‘substantial self-knowledge’ from

---

10 This is the position that Hurka, Best Things takes. He contends that there is no intrinsic value that attaches uniquely to self-knowledge:

This doesn’t mean self-knowledge has no value: understanding yourself remains as good as understanding other people or things. And sometimes it’s hugely beneficial, as when it helps you choose a rewarding career or improve your character. But its effects aside, self-knowledge doesn’t seem so special: it has neither the generality of scientific knowledge nor the extra importance of knowing your place in the world. While it’s good to know your inner states, it’s no tragedy if you don’t, and if being wrong about them helps you achieve other important good, that’s probably on balance a blessing (94).

Hurka restricts the domain of self-knowledge in a way that I do not: self-knowledge is knowledge pertaining specifically to oneself (and thus, one’s ‘inner states’), and not in your relation to the world. The latter form of knowledge, Hurka contends, does have a special value: while having correct beliefs about your relation to the world isn’t “a great intrinsic good”, having false beliefs “is a great evil” (90). This helps explain why there is something horrific about the idea that one has spent one’s life in an Experience Machine, but nothing especially uplifting about the knowledge that one has not. As I will explain in Chapter 3, knowledge of the relationships that one bears to the world can be self-knowledge: indeed, we can only explain the value we place on this knowledge by appeal to the intimate connection it has to our sense of self.
‘trivial self-knowledge’. Next, I’ll explain why we cannot account for the value of self-knowledge in purely instrumental terms. To do this, I will survey several different kinds of Instrumentalist Views: the Moral Self-Improvement Account, the Authenticity Account, and the Wellbeing Account. Having demonstrated that substantial self-knowledge has a value that goes beyond the instrumental, I will dedicate the rest of this Introduction to explaining why we ought to resist the urge to merely attribute it an intrinsic value. This will leave me with a challenge that I must meet over the next two chapters. Specifically, I will have to explain why self-knowledge can be valuable for its own sake while still providing a substantive account as to why it has this value.

Section 1: What Type of Self-Knowledge Is (Possibly) Worth Having for its Own Sake?

When I say that I endeavor to provide a non-instrumentalist defense of self-knowledge’s value, I do not mean to imply that every fact that we can possibly know about ourselves is non-instrumentally worth knowing. Indeed, I think that many of the things that we are capable of knowing about ourselves have at most instrumental value. To know whether I put the left sock on my foot before the right this morning is, in a sense, to know something about myself, but it would be hard to argue that my life is enriched in any meaningful way by virtue of having that little tidbit of self-knowledge in my possession. If my life is enriched by that knowledge at all, it will only be enriched insofar as that knowledge may help me achieve some further valuable end. Nor do I think that there is something inherently valuable about knowing exactly how far away my house is from New York City, the exact length of my small intestine, or whether my forth favorite type of cuisine is Greek or Thai.
In contrast, I do think that I would be better off, all things considered, if I were able to have accurate beliefs about where my talents really lie, whether I actually value what I purport to value, and why I actually made some of the major life decisions that I have made. These two classes of self-knowledge correspond to the distinction that Quassim Cassam has drawn between ‘trivial’ and ‘substantial’ self-knowledge, and it is to his account that I will now turn.

*Cassam’s Characterization of Substantial Self-Knowledge*

Before explaining what exactly differentiates substantial self-knowledge from trivial self-knowledge, Cassam first offers a list of the types of knowledge that intuitively belong to the former category. Substantial self-knowledge includes knowledge of our character, knowledge of our values, abilities, emotions and aptitudes, knowledge pertaining to the reasons that we have for holding certain beliefs about the world, and knowledge of what makes us happy. These types of self-knowledge, Cassam suggest, meet most (if not all) of ten conditions, which he describes as follows:

i. The Fallibility Condition: with substantial self-knowledge there is always the possibility of error. It’s not just a theoretical possibility that you are mistaken about, say, whether you are generous but an actual, real-life possibility. There isn’t even a presumption that you aren’t mistaken about such things because it might be a psychological fact about us humans that we are generally prone to thinking well of ourselves even if an objective view of the evidence would support a harsher assessment. It’s comforting to think that you are a generous

---

person even though your close friends can hardly fail to have noticed your
tendency to make yourself scarce when it’s your turn to buy the next round of
drinks.

ii. The Obstacle Condition: the possibility of error in such cases is a reflection of the
fact that for humans there are familiar and reasonably well-understood obstacles
to the acquisition of substantial self-knowledge. Such obstacles include
repression, self-deception, bias, and embarrassment. Some of us find it hard to be
honest with ourselves about our own limitations and that can make it hard to
acquire some types of substantial self-knowledge.

iii. The Self-Conception Condition: the existence of such obstacles to substantial self-
knowledge is related to the fact that, as Schwitzgebel puts it, this kind of
knowledge often ‘tangles with’ a person’s self-conception (Schwitzgebel 2012:
191). To know that you have a particular character you have to believe you have
that character, and it might be hard for you to believe that if your having that
character is at odds with your self-conception.

iv. The Challenge Condition: substantial self-knowledge can be challenged even in
normal circumstances. For example, if you assert that you have an aptitude for
dealing with difficult colleagues or that a charge of career would make you happy
there is room for the question, ‘Why do you think that?’, or for the retort ‘You
must be joking’. No doubt you have your reasons for thinking you have an
aptitude for dealing with difficult colleagues, but your reasons are not immune to
criticism and correction.
v. The Corrigibility Condition: substantial self-knowledge is corrigible, and its corrigibility is related to the fact that we are not necessarily authoritative about the matters to which such knowledge relates. You may not be in the best position to know about such matters and others might know better; your spouse may well have a much deeper insight into your character than you do.

vi. The Non-Transparency Condition: substantial self-knowledge can’t be got by employing the Transparency Method (TM). You can’t determine whether you are really happy with your new apartment by determining whether you ought rationally to be happy with it. Maybe you ought to be really happy with it but aren’t.

vii. The Evidence Condition: substantial self-knowledge is based on evidence. Many different kinds of evidence bear on substantial self-knowledge. If the question is whether you are in love with X, the evidence might include how you behave in the presence of X, how you feel in their presence (does your heart go pitter patter?) and what people who know you tell you about your state of mind. If the question is whether you are a racist the evidence will include how you behave towards, and think about, people who belong to racial groups different from your own.

viii. The Cognitive Effort Condition: acquiring substantial self-knowledge requires a degree of cognitive effort. For example, you don’t ‘just know’ your own character

---

12 Cassam takes this condition from Richard Moran, *Authority and Estrangement*, (Princeton: Princeton University Press, 2012). Moran argued that in order to know our belief about some proposition, \( p \), we needed to look not at our inner selves, but at the world around us to see what evidence there is in support of \( p \) (84). The question “Do you believe it will rain” is, on Moran’s account, transparent to the question “Will it rain?”.
or aptitudes; you have to work it out by reflecting on the behavioral evidence and on what other people tell you about you. The level of effort may be different in different cases but substantial self-knowledge can’t literally be effortless. If it were it would scarcely be ‘substantial’.

ix. The Indirectness Condition: in the terminology of Chapter 1, substantial self-knowledge is neither psychologically immediate nor epistemically immediate. It isn’t psychologically immediate because, as the Cognitive Effort Condition suggests, it requires reasoning or inference. It isn’t epistemically immediate because it depends on your having justification for believing other, supporting propositions. For example, you can’t know that you are generous if you aren’t justified in believing a range of supporting propositions about your actions. To know that you are generous it isn’t enough that you act generously; you also need to believe, and be justified in believing, that this is how you act.

x. The Value Condition: substantial self-knowledge matters in a practical or even a moral sense. If you are planning a trip to Spain it helps to know whether you can speak Spanish. As King Lear discovered, not knowing what will make you happy can result in your making bad choices, and we think some of forms of self-ignorance not just as cognitive but also as moral defects. Being unkind is bad in itself but made morally worse if it combined with the belief that one is kind.13

The items on Cassam’s list are not to be taken as necessary and sufficient conditions for substantial self-knowledge. Rather, they are meant to provide a rough indication of the sorts

---

13 Ibid., 31–33. The items on this list were reproduced exactly as they appeared in Cassam’s text. Note that, as he makes clear in Condition X, Cassam is an instrumentalist about the value of substantial self-knowledge.
of qualities that we associate with the self-knowledge that matters most. The difference between substantial and trivial self-knowledge, further, will be one of degree. “The more of these characteristics” that a kind of self-knowledge has, Cassam writes, “the more substantial it is.”

What Cassam has given us, in essence, is a cluster of properties that can be used as a rough heuristic to pick out facts about oneself that count as substantial self-knowledge. Not every piece of substantial self-knowledge will meet every one of these conditions. And some pieces of seemingly trivial self-knowledge, in contrast, will meet a startling number of these conditions. Consider Timothy, who believes he has an above-average tolerance for spicy food. Timothy doesn’t care about this, as he doesn’t have a strong opinion about the tastiness of spicy food. This fact about Timothy seems trivial—perhaps not as trivial as the fact that he is currently wearing socks, or the fact that the number of miles between him and Las Vegas is a prime number, but trivial nonetheless. And yet, it is a fact that meets most of Cassam’s conditions:

i. The Fallibility Condition: Timothy might have arrived his belief because he was mistaken about how tolerant other people are of spicy food, or because he was mistaken about how spicy the spicy food he’s eaten really was.

ii. The Obstacle Condition: there are obstacles to Timothy knowing this fact: even though he thinks that he doesn’t care about his above-average spice tolerance, he might be self-deceived about this fact. Perhaps he does care, but he doesn’t want to admit that he’s the sort of person who puts stock in such trivial things.

iii. The Challenge Condition: other people, upon hearing Timothy’s claim that he has an above-average tolerance for spiciness, could reasonably challenge the veracity of his claim.

iv. The Corrigibility Condition: other people watching Timothy eat spicy food might note, for instance, that he appears to be in more pain when eating spicy food than the average person.

v. The Non-Transparency Condition: Timothy, of course, can’t reason to the conclusion that he has an above-average tolerance for spicy food.

vi. The Evidence Condition: presumably, Timothy couldn’t reasonably draw a conclusion about his spice tolerance if he has only eaten spicy food once or twice, or if he has only eaten one type of spicy food.

vii. The Cognitive Effort Condition: Timothy would have had to reflect on his tolerance for spicy food, and his observations about other peoples’ tolerances, to conclude that he has an above average tolerance.

viii. The Indirectness Condition: for Timothy to have formed his belief, he must have reasoned from evidence to a conclusion. And he must have formed other beliefs along the way, such as the belief that the food he thought to be too spicy was, in fact, spicy.

ix. The Value Condition: Timothy’s knowledge will, at very least, help him in his practical deliberations about what sorts of new foods to try, and what to order at restaurants. It might also save him from embarrassment when eating spicy food with other people.
Indeed, Timothy’s spice tolerance only fails to meet one of Cassam’s conditions for substantial self-knowledge, *viz.* the ‘Self-Conception Condition’. And yet contra Cassam, we should not conclude from this that his knowledge is substantial. This is because it fails to meet this condition is all that we need to appeal to in order to understand why it fails as substantial self-knowledge. Even if it’s epistemically indirect, fallible, valuable, the result of cognitive effort and so forth, if it doesn’t speak to Timothy’s sense of himself, it isn’t substantial self-knowledge.

Now, the sorts of facts that *do* speak to our sense of self often do meet these other conditions. And they meet these conditions primarily because of the relation that they bear to our sense of self: our investment in being a certain way explains why we may be self-deceived about substantial facts about ourselves, why other people who care less may see these facts with more clarity than we do, and why we find these facts valuable. Likewise, given that substantial self-knowledge seems most typically to consist in facts that speak to our sense of self, it makes sense that these facts would rarely concern particular instances of our behavior, but would instead most involve the overarching conclusions that we’ve drawn from these instances. The fact that you made a charitable donation exactly three years ago says nothing substantial about your moral character. The fact that I am a dedicated philanthropist, which is itself a fact that I can derive only by looking at a wide range of my behaviors over time, does speak to my character and thus my sense of self.

*An Alternative: Substantial Self-Knowledge and Identification*

The Timothy example helps to illustrate the intuitive connection between substantial self-knowledge and our sense of who we are as individuals. Speaking colloquially, then, we
might say that substantial self-knowledge is knowledge about whatever it is that makes you you.

We can capture the rough contours of this category of self-knowledge by looking at it third-personally. Consider all of the facts that I might be able to learn about you. Presumably, the number of things that I could learn about you is infinite (if we include facts about Cambridge properties), but most of these facts will tell me nothing in particular about who you actually are and what you’re actually like. I can know that you have tulips planted in your front garden, and that you once drove through Detroit on a Sunday without being able to draw any reasonable conclusions about your character or values. At most, these facts give me a chance to speculate: I might jump, perhaps, from the fact that you have a garden to the supposition that you enjoy gardening. There are facts about you that leave me with even less room for speculation about what you’re like as an individual. Your blood type is AB+; you’re a prime number of miles away from Los Angeles; your left leg currently itches. These propositions may all be true, they do not give me sufficient grounds for making reasonable suppositions about what you might be like as an individual. In contrast, if I were to learn that you play in a Metallica cover band for fun on the weekends, that you’d rather miss your office’s Christmas party than your child’s Winter Recital, and that you have a tendency to be modest about your accomplishments, then I can on the basis of those facts begin forming a picture of what you’re like.

So far, this talk of ‘individual selves’ and ‘what makes us us’ has been vague and speculative. How might we distinguish facts about ourselves that speak to who we are as individuals from facts about ourselves that do not? For starters, I think we can say that the category of substantial self-knowledge, by and large, consists of facts pertaining to the ends
that we have set. To be a person is, according to a plausible rationalist picture of personhood at least, to be an end-setter.\(^{15}\) Who we are as individuals, in turn, is largely the result of the ends that we have set for ourselves; we have, over time, shaped ourselves by cultivating certain dispositions and ignoring others, embracing and rejecting different social identities, and setting particular goals and projects for ourselves. Your involvement in a Metallica cover band, and your commitment to your children, speak to the ends that you as an individual have set for yourself. In turn, it is understandable that you would care about knowing facts about the ends that you have set for yourself for non-instrumental reasons. After losing a Metallica cover band contest, for instance, you might find yourself wanting to know whether you lost because the judges had been bribed, or because your band isn’t actually as fantastic as you think it is. Likewise, you might question whether your devotion to your children isn’t in part a reaction to the lack of attention your own parents paid to you.

By tying the category of substantial self-knowledge to the ends that we set for ourselves, we are able to explain why not all knowledge of our talents is especially important to us. I haven’t set any ends for myself that require me to be a good singer, and so I have no special reason to find out just how bad a singer I really am. But Florence Foster Jenkins, whose ends and entire sense of self are bound up with her identity as a singer, \textit{does} have a special reason to find out the truth about her singing abilities: this knowledge, after all, is knowledge that directly pertains to one of the ends that she has set, and that speaks to her sense of who she is as an individual person.. Likewise, we are able to explain why biological facts about ourselves and facts about our Cambridge properties are not pieces of substantial

\(^{15}\) I will discuss this picture of personhood in detail in Chapter 1.
self-knowledge. Your blood type and the exact number of miles between you and Los Angeles presumably do not speak to any of the ends that you have set for yourself. You can, as such, find out facts about these topics without having learned anything that is especially relevant to who you are.

This distinction, however, cannot get us all the way to account of what types of propositions belong in the category of substantial self-knowledge. We constantly set ends for ourselves, and not all of the ends that we set speak to who we are in any meaningful sense. I chose to eat cereal rather than toast this morning, and I chose to do my laundry last Saturday. We can say that these behaviors reflect ends that I have set for myself: the end of eating cereal for breakfast, and the end of having clean clothing. But I could have easily chosen different ends for myself (toast for breakfast, laundry on Sunday) without being in any important sense a ‘different person’. The knowledge that I have about these ends, further, is clearly only instrumentally valuable: I care about knowing what I had for breakfast insofar as it informs next week’s grocery list, and I care about knowing when I last did laundry insofar as it helps me predict when I’ll have to do it next.

We cannot say, then, that substantial self-knowledge is any knowledge whatsoever that pertains to the ends that we set for ourselves. Many of the ends that we set for ourselves, after all, are mundane, habitual, and largely unrelated to our identity as individuals. And so, we must find some way of distinguishing the sorts of ends that are relevant to substantial self-knowledge from the sorts of ends that are irrelevant to it.

At this point, we might turn to Bernard Williams’ notion of ‘character’ to explain why not every piece of knowledge that we can have about our ends counts as substantial self-knowledge. To have a character, according to Williams, is to identify strongly with particular
‘ground projects’, such that one would experience a loss of those projects as a loss of meaning in one’s life.16 The ends that pertain to our ground projects and thus our character, which we might call ‘identity-constituting ends’, matter to us in a way that our other ends do not. And so too do we care about knowing information related to our identity-constituting ends in a way that we do not care about self-knowledge more generally.

While Williams’ discussion of the importance of ground projects might strike some as hyperbolic, I think he captures the intuitive idea that we most strongly with and affirm a specific subset of our ends. Much of the time, we set ends for ourselves out of habit, or uncritically. We don’t generally ask ourselves why we set these sorts of ends, both because we aren’t often particularly interested in the answer, and because we typically think that the answer is pretty obvious. Why do I do my laundry last Saturday? Because I needed to do laundry, and because I had time to do it on Saturday. I’m not ardently committed to doing laundry on Saturday, and I could have just as easily saved it until Sunday. In fact, I would suffer no great loss of self if I were to realize that I had actually done the laundry on Sunday, and not Saturday as I recall.

In contrast, the answers to questions like ‘Why did I stop talking to my best friend last year?’ or ‘Why am I so bothered by the fact that I’m no longer a vegetarian?’ are often genuinely unclear to us. I can question why I recently began eating meat after several years of vegetarianism—I can wonder, for instance, whether my current dietary preferences are

---

born out of a concern for my health, or a more basic desire to eat bacon again. Further, the answers to these types of questions matter to us. This is true even when we think that these answers are obvious. Even if I am fairly certain about why I started eating meat again, I would still think that the question was an important one for me to answer insofar as it pertains to one of the ends to which my identity was once strongly tied. If were to discover that I had been mistaken about my beliefs with regards to these ends, I would likely feel as though I’ve lost sight of who I actually am.

If substantial self-knowledge is self-knowledge that pertains directly\(^\text{17}\) to the ends that matter most to us as individuals, then the specific facts count as substantial self-knowledge will vary from one person to another. This is why I think that Florence has a reason to care about facts pertaining to her musicality in a way that I do not. Facts about her musicality are substantial facts about Florence precisely because they speak to one of her identity-constituting ends. Facts about my lack of musicality are not substantial facts about me precisely because they do not.

Given this, it will be difficult to make determinations about whether a specific fact about an individual counts as substantial absent sufficient information about the background circumstances of that individual’s life. To make that determination, we would need to know something about what ends that individual has set for herself, and how she has come to identify with the ends that she has set. We would need also to know how that fact relates, or

---

\(^{17}\) There will be some vagueness about what it means for information to pertain ‘directly’ to one of these ends. Certainly, there is some information about our identity-constituting ends that we don’t care about: a musician might want to know whether he’s actually talented at his chosen instrument, but not whether he picks up that instrument more often with his right or left hand. The former fact matters directly to the ends that he has set for himself as a musician in a way that the latter does not. A full description of what it takes for information to pertain directly to an identity-constituting end is beyond the scope of this project.
fails to relate, to her identity-constituting ends. Outside of particular cases, the most we can reasonably do to delineate substantial self-knowledge from trivial self-knowledge is to point to broad categories of facts, such as facts about our values, talents, significant life events and so forth, insofar as these, insofar as these categories often closely track the types of identity-constituting ends that agents set for themselves.

Section 2: The Instrumental Value of Substantial Self-Knowledge

Even if it is clear why we value some types of self-knowledge over others, it is not yet clear why we do this or whether we ought to do this. Most typically, philosophers have answered questions about self-knowledge’s value by offering some iteration of what I will call ‘The Instrumentalist View’. The Instrumentalist View holds that self-knowledge is valuable insofar as it is a means to some independently valuable end. The view can be taken to be either a descriptive claim about how we actually value self-knowledge, or (as it is most typically interpreted) as a normative claim about how we should value self-knowledge.

The Instrumentalist View comes in various flavors, depending on which independently valuable end it associates with self-knowledge. Before moving on to my critique of the view, I’ll briefly summarize three common formulations of the view: the Moral Self-Improvement Account, the Authenticity Account, and the Well-Being Account.
The Moral Self-Improvement Account

Historically, the value of self-knowledge has been tied to with the moral disvalue of self-deception. For Aristotle, Plato, Kant, Butler, Smith and others, self-deception was the root of many (if not all) human evils. Self-deception, Butler warns:

…is unfairness; it is dishonesty; it is falseness of heart: and it is therefore so far from extenuating guilt, that it is in itself the greatest of all guilt in proportion to the degree it prevails…

Self-deception allows us to ignore our moral failings, thus halting moral self-improvement, and giving us the unique power to commit immoral actions without having to bear the burden of a guilty conscience. We find an echo of Bishop Butler in Adam Smith’s condemnation of self-deception. Self-deception, he writes, is the “fatal weakness of mankind” insofar as it allows us to “persevere in injustice.” Kant, meanwhile, positioned a duty of self-knowledge at the center of his moral framework. The ‘First Command’ of all self-regarding duties, was to:

…know (scrutinize, fathom) yourself,” not in terms of your natural perfection (your fitness or unfitness for all sorts of discretionary or even commanded ends) but rather in terms of your moral perfection in relation to your duty. That is, know your heart—whether it is good or evil, whether the source of your actions is pure or impure, and what can be imputed to you as belonging originally to the substance of a human being or as derived (acquired or developed) and belonging to your moral condition.

18 Aristotle, Nichomachean Ethics; Butler, Sermons; Kant, Metaphysics of Morals; Smith, Moral Sentiments.

19 Butler, Sermons, 120.

20 Ibid.

21 Smith, Moral Sentiments, III.I.91-92.

22 Kant, Metaphysics of Morals, 6:441.
Fulfilling this duty, Kant suggested, will help to rid us of our self-conceited tendency to overinflate the goodness of our moral actions and characters. We cannot, in other words, fruitfully seek to fulfill our duties, if we are self-deceived about our inborn tendency to stray from the commands of morality.\(^23\)

*The Authenticity Account*

More recently, philosophers have located the value of self-knowledge in its connection to authenticity. There is philosophical disagreement both about what it means to be ‘authentic’ about what types of self-knowledge, if any, are required for authenticity, and about whether authenticity is itself instrumentally or non-instrumentally valuable.\(^24\) At one end of the ‘authenticity spectrum’, we find Jean Paul Sartre, who understood the connection between authenticity and self-knowledge rather negatively: to have authenticity was not so much to know oneself as it was to lack a certain type of self-deception, which Sartre called ‘bad faith’.\(^25\) The person with bad faith fails to grasp something about her nature: she sees herself purely as an ‘objective facticity’, that is, as a being with determinate characteristics over which she had no control, while denying her status as a self-determining entity capable of transforming herself and her character. To overcome bad faith, and thus to achieve


authenticity, is to recognize the duality of one’s nature as both an objective facticity and a transcedently free self-creator.\textsuperscript{26} Thus, insofar as authenticity requires self-knowledge, the self-knowledge that it requires is knowledge not about oneself as an individual, but rather about the duality of one’s fundamental nature.

At the other end of the spectrum, you will find more substantive accounts of the ideal of authenticity. For philosophers like Lionel Trilling, J. David Velleman, and Michael Lynch, living authentically was a matter of embodying one’s true self by living in accordance with one’s actual desires. What it means to have a ‘true self’ or what it means for a desire to ‘actually’ be along to a person is subject to philosophical debate. Following, Velleman, we might capture this version of the ideal by looking at a paradigmatic example of an inauthentic person:

This person laughs at what he thinks he is supposed to find amusing, shows concern for what he thinks he is supposed to care about, and in general conforms himself to the demands and expectations of others. The motives that his behavior is designed to simulate are motives that he doesn’t genuinely have.\textsuperscript{27} Velleman’s paragon of inauthenticity, in other words, is a conformist. The authentic person, in contrast, lives in a way that accurately represents who and what he truly is.\textsuperscript{28} Michael Lynch offers a more idealized version of this view. For him, the authentic person is someone

\begin{itemize}
\item \textbf{28} J. David Velleman, \textit{How We Get Along}, (New York: Cambridge University Press, 2009), 102. Velleman makes clear that inauthenticity need not be conscious or deliberate, but rather can be the result of self-deception (90–92).
\end{itemize}
who identifies with the desires that guide his actions. He is, in other words, not simply being himself, but being a version of himself that he can reflectively endorse.29

Any account of self-knowledge that locates its value in its connection to authenticity will be subject to two potential lines of criticism. First, if authenticity turns out to not itself be a valuable ideal, then we will not be able to use it to ground the value of self-knowledge. And second, if we could demonstrate that self-knowledge is not to be required for authenticity, then we will not be able use its purported connection to authenticity to explain its value.

In support of the first line of criticism, Christopher Lasch has suggested that authenticity is problematically narcissistic,30 Charles Taylor have both criticized the ‘culture of authenticity’ as unduly individualistic, relativistic, and devoid of meaning,31 and Feldman has argued that authenticity cannot be justified by appeal to morality, rationality, or well-being.32

The second criticism has been pursued most forcefully by Hazlett and Feldman, who argue that authenticity, however we describe it, does not require self-knowledge, and is in

---


32 Feldman, *Against Authenticity*. 
fact sometimes inconsistent with self-knowledge.\(^{33}\) If we understand authenticity in Sartrean terms, they contend, then we end up with a circular account of self-knowledge’s value insofar as authenticity is essentially a species of self-knowledge. Understanding authenticity as a matter of ‘being true to yourself’ or ‘wholeheartedly identifying with your effective desires’\(^{34}\) gets us at most a weak connection between authenticity and self-knowledge: while knowing ourselves may sometimes help us act authentically, it is not required for authentic action. We can, in other words, act upon the desires we have, or the desires we endorse, without knowing exactly what those desires are.\(^{35}\)

*The Well-Being Account*

Persuaded by both Feldman and Hazlett’s arguments\(^{36}\), Cassam has argued that self-knowledge is valuable insofar as it positively contributes to our well-being. The value of well-being, in turn, needs no argument: it is self-evidently intrinsically valuable. Although Cassam never fully explicates how he understands the term ‘well-being’, he seems to be relying on either a utilitarian or desire-satisfaction model. As he describes it,

---

\(^{33}\) Hazlett and Feldman, “Authenticity and Self-Knowledge.”

\(^{34}\) See Lynch, *True to Life* for a discussion of this latter view. Lynch argues that being true to oneself consists not simply in acting upon one’s desires, but specifically in acting on those desires that one reflectively endorses (125).


Knowing what you are like, good at, what makes you happy, what is important to you, or how you feel can improve your choices and thinking in ways that objectively make your life go better.37

The plausibility of Cassam’s claim cannot be denied: although there is some empirical evidence suggesting that minor self-deceptions can at times increase our well-being, it would be hard to doubt that our lives go better when we are able to accurately form beliefs about our desires, talents, experiences, and attitudes.38

Section 3: Against the Explanatory Exhaustiveness of the Instrumentalist View

At the outset, I must make clear that I do not take issue with the Instrumentalist View as an explanation of self-knowledge’s value. Certainly, self-knowledge is valuable as a means to all sorts of other instrumentally and non-instrumentally valuable ends. My issue with the Instrumentalist View only arises if one takes it to be an exhaustive explanation of self-knowledge’s value. I will grant, in essence, that the Instrumentalist View has something to say about self-knowledge’s value, while denying that it has the last word on the subject. Given this, rather than critique each account separately, I will focus on one type of the Instrumentalist View—Cassam’s Well-Being account—to show why it cannot fully describe the value many of us place, and think we ought to place, in self-knowledge.

The Instrumentalist View, as I have said, can be taken as either a descriptive thesis about how we value self-knowledge, or a normative thesis about how we ought to value self-

---

37 Ibid, p. 225. Additional evidence for this interpretation can be found at p. 277, when Cassam suggests that the precise ways in which self-knowledge contribute to well-being are a matter of empirical, rather than philosophical investigation.

knowledge. To see where a descriptive reading of Cassam’s version of the Instrumentalist View falls short, consider the following hypothetical. Suppose that Cassam tells you that self-knowledge is instrumentally valuable only insofar as it positively contributes to your overall wellbeing. Two results fall out of Cassam’s claim: first, that self-knowledge will cease to be valuable if it turns out to be irrelevant or deleterious to wellbeing, and second, that the reasons that we have to seek self-knowledge will ultimately boil down to the reasons that we have for wanting well-being.

So far so good. Now, suppose that I interrupt your chat with Cassam to inform you that I have discovered something about you that you don’t know. If you were to find out this piece of information, I warn you, it would radically change your conception of who you are and how your life is going. This secret is, in other words, that might be revealed on an episode of the Jerry Springer Show. Suppose further that I know you better than anyone else in the world, and that I have every reason to believe that nothing good will come from you knowing this piece of information about yourself—it won’t guide your future deliberations, or lead you to a sense of inner peace, and it may even cause you emotional harm.

Would you want me to tell you the secret? In my experience, people part ways on their answer to this question. Some emphatically want to know, some empathically want not to know, and others are genuinely unsure of whether or not they want to know. This is enough to render the descriptive reading suspect. But does it leave us with any reason to reject the more plausible normative reading? Consider what can be said about the rationality of the different answers that one might give to my offer. Which answer is the right one? I want to suggest that these answers are all rationally on a par. That is, we wouldn’t think that there is anything amiss about someone who strongly prefers one alternative to the other. That
is, unless we assume that rationality always demands that we act in ways that aim to maximize utilitarian wellbeing, then I think each of these choices will appear to most people as intelligible and ceteris paribus reasonable. If this is right, then it gives us reason to doubt that the Instrumentalist View is capable of offering an exhaustive explanation of the value that we can reasonably place in self-knowledge. If self-knowledge were only valuable instrumentally, then it would be unreasonable to want to know something that would undermine the very state that gives it valuable. This is clearly the case with other instrumentally valuable things. If Midas had known beforehand that having the ability to turn everything to gold would cause him nothing but strife, it would have been unreasonable for him to ask Dionysus for that power.

Someone sympathetic to Cassam’s account might object that I have rigged this thought experiment in my favor. After all, the very act of telling you that I know a shocking secret about your life could itself negatively impact your well-being. Even if you choose not to find out the secret, you might still spend sleepless nights wondering about what I would have told you. For some, this state of knowing-that-you-don’t-know could in fact be worse than the psychological damages that the secret would have caused. But we can tweak the example, to account for this wrinkle. Suppose I promised to give you a dose of rohypnol that would erase your memory of our conversation if you chose to turn down my offer. Even if you were able to turn down my offer without any ill effects, I think it would still be intelligible for you to choose either to find out or not find what I know about you.

39 There might be situations in which one or more of these answers is unreasonable. If you have a history of mental instability, and obligations to remain mentally well for the sake of your child’s welfare, then it may be unreasonable for you to choose a course of action that heightens your chances of being able to fulfill that obligation.
If this hypothetical is unconvincing, consider another. Suppose that, after much research, psychologists found out that, while self-knowledge typically increases levels of well-being by 10%, there is another epistemic state—call it ‘minor self-overestimation’—that on average increases well-being by 15%. By ‘minor self-overestimation’, I have in mind the state of having beliefs about oneself that are, to a minor degree, overly positive given the evidence available. A person in this state would not have a radically false self-conception. Rather, they would conceive of themselves in a way that is just a little too good to be true.40 If self-knowledge is only good insofar as it positively contributes to well-being, then minor self-overestimation would ceteris paribus be better a better epistemic than self-knowledge. If you could choose between the two without any ill effects, then you would have more reason to choose the latter than the former. Indeed, according to the expected utility account of rationality, it would be patently irrational for you to choose self-knowledge to minor self-overestimation. This result, at very least, represents a significant departure from the way that we value these two epistemic states.

The Instrumentalist View: Conclusion

Similar objections can be raised towards both the Moral Self-Improvement Account and the Authenticity Account. We can still intelligibly value self-knowledge that is irrelevant or even deleterious to our moral progress. Likewise, even if self-knowledge turns out to be inessential to authenticity, as Feldman and Hazlett argue, we can still cogently understand

why someone would prefer a life that has both self-knowledge and authenticity to a life that has only authenticity.

My arguments against the Instrumentalist View have, of course, involved blatant intuition pumping. I take it, however, that these intuitions are widely shared. People commonly try to uncover unpleasant truths about themselves, and it is doubtful that their reasons for seeking these truths are always purely instrumental, or that our judgments about the reasonableness of these sorts of inquiries are responsive purely to the effect these discoveries will have on their wellbeing. Further, we often praise people who are able to face themselves as they really are, and think less of people who don’t. I have not yet defended these practices, but I think that their prevalence places a significant burden on the Instrumentalist View. If we take the view to be a purely descriptive theory of how we value self-knowledge, then it fails in its task. As a matter of psychological fact, many of us do not care about self-knowledge for purely instrumental reasons. But as a normative theory, the view also falls short, insofar as it requires us to revise many of our pre-philosophical judgments about rationality, virtue, vice and the role that an accurate self-conception plays in a good life.

An instrumentalist could push back here. The aim of philosophy should not be to salvage everything that people happen to believe. Rather, philosophy should drive us to scrutinize what we believe, and reject those beliefs that do not stand up to critical inquiry. And so, if the Instrumentalist View requires us to revise some of our pre-philosophical judgments about what is reasonable and virtuous, then so much the worse for those judgments. This presents those of us who are unsympathetic to the Instrumentalist View with
a formidable task: if we think that these particular judgments are worth rescuing, then we must figure out some way of rescuing them.

Section 4: The Ineffability of Intrinsic Value

If self-knowledge’s value cannot fully be captured in instrumental terms, how else can it be captured? One move that could be made here is to ascribe an intrinsic value to self-knowledge: self-knowledge would, as such, innately possess a value that ‘stands alone’ and that does not rely on any connections to other valuable entities or states of affairs.41

This is the move that Eric Schwitzgebel has made in his commentary on Cassam’s *Well-Being for Humans.*42 In defense of his claim that self-knowledge is intrinsically valuable, Schwitzgebel provides three arguments. The first argument, which he calls “The Argument from Addition and Subtraction” relies on the intuitive idea that something is lost when we subtract self-knowledge from the world, even if everything else in the world is kept constant, and that something is likewise gained when we add self-knowledge to it. If we imagine that we lack some of the self-knowledge about our character or biases that we presently have, for instance, but that no negative consequences follow from this lack, we can still think that we’ve lost something in the subtraction. This suggestions, *contra* Cassam, that

41 A similar characterization of self-knowledge’s value can be found in James Griffin’s *Well-Being: Its Meaning, Measurement and Moral Importance* (Oxford: Oxford University Press, 2002), when he remarks that “[s]imply knowing about oneself and one’s world is part of a good life. We value, not as an instrument but for itself, being in touch with reality, being free from muddle, ignorance, and mistake.” (67).

there’s something valuable about self-knowledge that can’t be cached out in terms of the effects it can have on our lives.\(^\text{43}\)

Schwitzgebel’s second argument is an argument from nearby cases. If we assume that knowledge in general is intrinsically valuable, and more specifically that knowledge of other people is intrinsically valuable, and if we grant that self-knowledge is relevantly similar to interpersonal knowledge, then we ought to also grant that self-knowledge is intrinsically valuable.\(^\text{44}\)

To insist otherwise would be incredibly odd: “It would,” Schwitzgebel suggests, “be weird if psychological knowledge of your friends and family had intrinsic value but psychological knowledge of yourself did not.”\(^\text{45}\)

Schwitzgebel’s third argument, “The Argument from Identity”, contends that knowing a psychological fact about yourself is identical to possessing a set of dispositions or capacities with respect to your own psychology. To know that you are an extrovert, for example:

…is in part the capacity to say, truly and sincerely, ‘Yeah, I’m an extrovert’. It is in part the capacity to respond appropriately to party invitations, by accepting them in anticipation of the good time you’ll have. It is in par not to be surprised to find yourself smiling and laughing in the crowd.\(^\text{46}\)

\(^{43}\) Ibid, 2-3.

\(^{44}\) This second argument is similar to the argument that I will ultimately be endorsing in Chapter 2. Nevertheless, my position on the non-instrumental value of self-knowledge is importantly distinct from Schwitzgebel’s: while Schwitzgebel is focused on attributing intrinsic value to self-knowledge, I’m ultimately making a different claim about the non-instrumental reasons we have to care about knowing ourselves. By paying attention to the reasons that we have to know ourselves, I suggest, we can make sense of why not all self-knowledge is worth having in the way that Schwitzgebel suggests.

\(^{45}\) Ibid, 4.

\(^{46}\) Ibid, 5.
These capacities, according to Schwitzgebel, are intrinsically valuable insofar as they are central to living a meaningful life. Without them, our lives would be impoverished, even if everything else were somehow to remain fixed.

Schwitzgebel grants that his arguments rely on intuitions, and more specifically on his intuitions. Thus, they at best can be used to reveal the fact that he regards self-knowledge to have intrinsic value. But this limitation, Schwitzgebel contends, reflects a more general methodological limitation of any sort of inquiry into values: we can, at best, explain to others what we value, and invite them to share in those values. Still, we must ask whether Schwitzgebel, by appealing to a bare notion of intrinsic value, is able to provide a persuasive invitation. His second argument, for instance, locates self-knowledge’s intrinsic value via its inclusion in a broader category (the category of ‘knowledge’) that is itself assumed to be intrinsically valuable. But if this is so, then we might question why we ought to care in particular about knowing ourselves, rather than knowing about facts about other people or the universe at large. Our special interest in self-knowledge could not be justified via its appeal to intrinsic value, as its intrinsic value in no way distinguishes it from all the other intrinsically valuable knowledge that we could have learned in the time we spent trying to discover ourselves. Rather, if we grant that self-knowledge is intrinsically valuable because it is a type of knowledge, then it would seem that we would have to resort to some sort of

---

47 Ibid, 3. Schwitzgebel raises and responds to this concern with regards to the first argument, but it applies to the second and third arguments as well.
Instrumentalist Account to explain why we have special reasons to know ourselves that we don’t have to know most other things.\(^{48}\)

This objection hints at a larger issue that will arise with any sort of account that attempts to explain why we are worth knowing for non-instrumental reasons via an appeal to self-knowledge’s intrinsic value. Intrinsic values are, as G. E. Moore’s Open Question argument demonstrates, conversation-enders.\(^{49}\) If you ask me why happiness was intrinsically valuable, I could do no more than tautologically inform you that ‘happiness makes us happy—and that’s just a good thing.’ This means that intrinsic values, by their very nature, do little to explain why we come to care about certain things non-instrumentally. Instead of moving the debate about self-knowledge’s value forward, an appeal to intrinsic value simply puts a halt to it. Further, attaching such a value to self-knowledge would leave us with additional questions. Why, for instance, is it the case that only certain types of self-knowledge seem worth having for their own sake? Likewise, to echo Schwitzgebel, if there is something intrinsically valuable about knowing oneself, we should also think that there is intrinsic value in knowing other people. But generally speaking, we find ourselves with wanting to know ourselves in a way that we do not want to know most other people.\(^{50}\)

---

\(^{48}\) This point has also been made by Michael Cholbi in his unpublished paper “The Value of Self-Knowledge” 14.

\(^{49}\) G.E. Moore, *Principia Ethica* (Cambridge: Cambridge University Press, 1993). This feature of intrinsic value is what prompts Cassam, *Self-Knowledge* to reject the possibility that substantial self-knowledge is intrinsically valuable. “If Y is valuable but it is impossible or inappropriate to attempt to explain what makes it valuable in more basic terms, then Y has intrinsic value,” Cassam observes. But since we *can* reasonably ask why self-knowledge is valuable, he concludes, it must not be intrinsically valuable (212).

\(^{50}\) The exception, of course, is the people with whom we share close relationships. We often value knowledge of our loved ones for its own sake, and not for its instrumental value. It is telling that, in presenting his Argument from Nearby Cases, Schwitzgebel focuses on the interpersonal knowledge that we can have of our friends, rather than interpersonal knowledge more generally. This suggests that what is ‘doing the work’ in this
may think that there is intrinsic value in knowing whether your self-perception accurately reflects reality, but it would be odd if you thought that it was intrinsically valuable for you to know this fact about the stranger standing next to you at the bus stop. Even if we grant that self-knowledge has intrinsic value, we cannot use it to explain why we seem to have special reasons to know ourselves that go far beyond the reasons that we have to know people more generally.

Finally, appeals to the intrinsic value of self-knowledge cannot settle the normative question of whether we ought to know ourselves. Many things—pleasure, health, wisdom, justice, the natural world and freedom, to name a few—might be said to have intrinsic value. But this does not mean that we are under any sort of obligation to engage with these things. It is consistent to believe that the natural world is intrinsically valuable and to have no desire to ever go on a hike, watch the sun set, or smell a rose. At most, we might think that nature’s intrinsic value at most gives us a reason not to disrespect or destroy it.

There is another move that can be made here. Instead of considering what value-properties self-knowledge may or may not have, we can instead focus on what reasons that we have to value self-knowledge non-instrumentally, or for its own sake. By focusing on how we value self-knowledge, and why we value self-knowledge, rather than on what type of value self-knowledge does or does not have, we will be able to more fruitfully engage with the question of why self-knowledge matters to us in a way that knowledge more generally

---

argument is not the bare intrinsic value of interpersonal knowledge, but rather the special value that we attach to knowledge of our loved ones. What this special is, and where it comes from, will be discussed in Chapter 2.

51 I will be using ‘non-instrumentally valuable’ and ‘valuable for its own sake’ interchangeably.
does not. While the question of what value self-knowledge has or does not have is a metaphysical question about one of its purported properties, the questions about how and why we value self-knowledge in the particular way that we do are normative questions that invite us to consider the reasons that we have for caring about self-knowledge in this distinctive way, and the role that self-knowledge plays in our broader conception of how we ought to treat ourselves and live our lives.

Section 5: The Structure of This Dissertation

In this dissertation, I will defend the practice of valuing self-knowledge for its own sake by locating a non-instrumental reason that we all have to know ourselves. In Chapter 1, I will offer a close examination of the concept of self-respect. In doing so, I will argue that, just as we ought to respect ourselves as practically rational agents, so too ought we respect ourselves as epistemically rational agents. To do this requires that we hold ourselves accountable to sufficiently stringent standards of belief formation, and also that we be especially concerned with the beliefs that we form about the ends that matter most to us.

Epistemic self-respect, however, only gets me part of the way towards my claim about the non-instrumental reasons that we have to know ourselves. While it can be used to explain why we ought to care about how we form beliefs, and why we should be especially concerned with beliefs relating to those ends with which we most identify, it cannot itself tell us what particular ends we ought to set for ourselves. Thus, while considerations of epistemic self-respect can explain why we ought to be concerned about how we pursue a project of knowing ourselves, it cannot provide us with an explanation as to why that project is one that we ought to have.
In Chapter 2, I argue that self-love can explain why we ought to have a non-instrumental commitment to self-knowledge. This argument begins with a close examination of the normative requirements of loving relationships. To be in a loving relationship with another person is, in part requires a commitment to loving that person. And loving someone gives us a non-instrumental reason to know that person. And so, among the normative requirements of loving relationships is a requirement to be committed to knowing the people with whom we share them for the sake of knowing them. I then argue that we are in a relationship with ourselves that is answerable to normative requirements of both self-love and self-respect. Self-love generates non-instrumental reasons to know ourselves, while self-respect demands that we take seriously those reasons by pursuing our project of self-inquiry in accordance with suitably stringent epistemic standards. We are therefore morally required not just to think about who we are, but to think about who we are clearly and carefully enough so that we may come to know ourselves.

Finally, in Chapter 3, I explain why we need not appeal to the negative consequences that often follow from being self-deceived in order to explain why self-deception is an undesirable state to be in. Rather, as I argue, self-deception is in and of itself objectionable insofar as it constitutes a ‘perversion of valuing’. By this, I mean that it simultaneously reflects what ends we most value, while also impeding our ability to ever achieve or properly engage with those ends. In this way, our self-deception represents not simply a failure to hold ourselves accountable to appropriate epistemic standards when forming beliefs, but also a failure to do so with regard to one of the ends that matters most to us, and that is thus most demanding of our epistemic attention. I will argue, however, that not all cases of self-
deception have this feature—some, in fact, can be part of what it takes for us to engage appropriately the ends that we set for ourselves.
CHAPTER 1: SELF-KNOWLEDGE AND EPISTEMIC SELF-RESPECT

Introduction

In her essay “On Self-Respect,” Joan Didion described how she initially came to respect herself:

I had not been elected to Phi Beta Kappa. This failure could scarcely have been more predictable or less ambiguous (I simply did not have the grades), but I was unnerved by it; I had somehow thought myself a kind of academic Raskolnikov, curiously exempt from the cause-effect relationships which hampered. Although even the humorless nineteen-year-old that I was must have recognized that the situation lacked real tragic stature, the day that I did not make Phi Beta Kappa nonetheless marked the end of something, and innocence may well have been the word for it. I lost the conviction that lights would always turn green for me, the pleasant certainty that those rather passive virtues which had won me approval as a child automatically guaranteed me not only Phi beta Kappa keys but happiness, honor and the love of a good man…To such doubtful amulets had my self-respect been pinned, and I faced myself that day with the nonplussed apprehension of someone who has come across a vampire with no crucifix at hand.52

To have self-respect, Didion suggests, is in part to see oneself clearly. But what does it mean to see oneself clearly in this context? And what is the nature of the relationship that obtains between self-respect and self-knowledge?

This chapter seeks to answer both of these questions via an extended exploration of what it means to respect oneself as a person. In it, I argue that self-respect requires that we appreciate ourselves as both practically rational and epistemically rational beings. It is this

second form of self-respect, which I call ‘epistemic self-respect,’ that will ultimately help explain in the next chapter why we ought to be non-instrumentally invested in knowing ourselves.

The structure of this chapter is as follows. First, I discuss the type of respect that I have in mind. This respect, which Darwall calls ‘recognition respect,’ is the type of respect that we are commonly assumed to owe to ourselves and others simply in virtue of being persons. Next, I consider what parts of personhood are themselves proper objects of respect. Although rationalists have traditionally assumed that it is our status as practically rational agents that make us deserving of respect, I explain why our reasons for seeing practical rationality in this way apply equally well to epistemic rationality. That is, if practical rationality makes us deserving of respect, then so too must epistemic rationality. In Sections 3 and 4, I consider what this means for our understanding of the interpersonal and intrapersonal requirements of self-respect, respectively. I argue that respecting ourselves as rational agents requires us to hold ourselves accountable to sufficiently stringent standards of epistemic reason, especially in relation to our most important life projects and ends. However, as I make clear in the conclusion to this chapter, we cannot simply appeal to the requirements of epistemic self-respect to explain why we ought to set self-knowledge as one of these ends.

**Section 1: Respect**

When Dr. Seuss wrote that “people are people, no matter how small,” Dr. Seuss, *Horton Hears a Who!* (Random House, 1954).
persons. This form of respect, which Darwall calls “recognition respect”\textsuperscript{54} lies at the foundation of our moral lives. When others fail to treat us with respect, they become the proper objects of moral censure. When we fail to treat ourselves with respect, we open ourselves up to well-intentioned criticism from those around us.

But the claim that all persons deserve respect \textit{qua} persons is, by itself, as unilluminating as it is uncontroversial. What does it mean to respect a person? And why are persons, rather than chairs or rocks, afforded this unique moral standing? To begin, we’ll look more closely at the general phenomenon of recognition respect, and then move on to the specific case of recognition respect for persons.

For Darwall, the phenomenon of recognition respect goes far beyond recognition respect for persons. As he describes it, to have recognition respect for something, be it a person or some other recognition respect-worthy entity, “consists, most generally, in a disposition to weigh appropriately in one’s deliberations some features of the thing in question and to act accordingly.”\textsuperscript{55} We can have recognition respect for the law, for the Office of the President, and for nature, among other things. Some types of recognition respect are non-moralized: they simply consist in assigning appropriate weight to some entity in one’s deliberations about how to act. At other times, recognition-respect can be “essentially a moral attitude” such that an “inappropriate consideration or weighing of that

\footnote{\textsuperscript{54}Stephen J. Darwall, ‘Two Kinds of Respect,’ \textit{Ethics} \textbf{88} (1977): 36-49. Throughout this dissertation, I will be using the terms ‘respect’ and ‘recognition respect’ interchangeably.}

\footnote{\textsuperscript{55}Ibid, 38.}
fact [that something is deserving recognition respect] would result in behavior that is morally wrong.”

That these facts, whatever they may be, deserve weight in our deliberations, is not accidental, but rather arises out of the nature of the facts themselves. To have recognition respect for the law, for instance, “is to be disposed to regard to regard the fact that something is the law as restricting the class of actions that would be morally permissible. Having recognition respect for the law, in other words, in part requires that we appreciate that the law, by its very nature, restricts the range of actions available to us.

Further, to have recognition respect for something requires that we factor it into our deliberations, and be willing to moderate our actions towards it, for the right sorts of reasons:

To have recognition respect for something is to regard the fact as itself placing restrictions on what is permissible for one to do. It is of course true that one can “be respectful” of something without having any respect for it…This will be the case if one behaves as one who does have respect would have behaved, but out of motives other than respect. For example, a person participating in a legal proceeding who in fact has no respect for the judge (i.e., for the position he occupies) may take great pains to be respectful in order to avoid a citation for contempt. Such a person will restrict his behavior toward the judge in ways appropriate to the role that he plays. But his reason for so doing is not that the mere fact of being the judge is itself deserving of consideration, but that the possibility of a contempt citation calls for caution.

And so too is proper regard essential to recognition respect for persons. An example can bring this point out. Suppose a white person acts respectfully towards every black man whom she meets not because she thinks that they deserve respect because they are persons, but

---

56 Ibid, 40.
57 Ibid, 40.
58 Ibid, 40.
59 Ibid, 40-41.
instead because she believes that that African American males are prone to violence, and fears that they will harm her if she fails to act respectfully towards them. This woman does not have a deficient type of recognition respect for African American males; instead, she has no recognition respect for them at all.

We can summarize the requirements of recognition respect for persons as follows. To have recognition respect for some person, X, one must have:

1. The judgment that X is worthy of recognition respect by virtue of belonging to a type, ‘person’, that deserves recognition respect, and;
2. A willingness to properly constrain her behavior towards X on the basis of that judgment, where
   a. What counts as proper constraint will be determined by facts about what it is to be a person.

While Darwall leaves vague exactly how recognition respect for persons requires us to constrain our behavior, a quick survey of the types of behaviors that we commonly deem ‘disrespectful’ should give us some idea of what these constraints generally look like. Your teenage son might charge you with disrespecting him if you insist on enrolling him in karate instead of community theatre, while your date may call you out for being disrespectful if you take it upon yourself to order for her at a restaurant. And the library patrons could certainly complain about your disrespectful behavior if you insisted on loudly Skyping with your Aunt Cathy in the silent reading room.

---

60 What recognition respect requires, Darwall observes “is not a matter of general agreement, for this is just the question of what our moral obligations or duties to other persons consist in” (“Respect,” p. 38).
These cases should also give us some idea of what it is about personhood that makes it (and thus persons) inherently deserving of respect. When we ask why these cases are instances of disrespect, we are met with a familiar rationalist picture of personhood. To be a person is to be the sort of entity capable of making choices for itself. To respect someone as a person, then, often requires us to respect the choices that she makes for herself. You should respect your son enough to let him sing show tunes, your date enough to let her order for herself, and the library patrons enough to let them read in silence.

That respecting a person often requires that we not interfere unduly with the choices that other people make for themselves helps us explain why even well-intentioned paternalism is so often considered disrespectful. As Seana Shiffrin observes:

> Paternalistic behavior is special because it represents a positive (although often sometimes unconscious or sometimes caring) effort by another person to insert her will and have it exert control merely because of its (perhaps only alleged) superiority. As such, it directly expresses insufficient respect for the underlying valuable capacities, powers, and entitlements of the autonomous agent.  

What else does respect require of us? Presumably, it demands that we not treat others in ways that undermine their very ability to act autonomously, or that are dehumanizing. Dehumanizing behavior, as the name suggests, often aims at destroying the very capacities of an agent that makes her a person. When I torture a terrorist in order to get information out of them, I do so with the intention of taking away their very ability to choose what information they do and do not share with me. Of course, not all dehumanizing actions are performed with some other end in mind—rather, we sometimes dehumanize others for the simple sake of dehumanizing them. But the very notion that there are some actions that are, in and of

---

themselves, dehumanizing reflects the rationalist thought that there are some things to which no person could rationally consent.

Some of the behaviors that we associate with ‘respect’ have a more symbolic connection to personhood. In my culture, there is nothing rude about wearing shoes inside. To my Indian friend’s mother, however, it would be the height of disrespect to not take one’s shoes off at the door. When I enter her home, I make a point of following that custom, and in doing so I communicate that I respect her enough to obey the etiquette norms that matter to her. Once again, we can tell a story about the value of persons. My friend’s mother has chosen to adopt certain cultural norms as her own. I may not share those norms with her, but my respect for her and her ends requires that I abide by them when in her home.62

The requirements of self-respect are more robust than the requirements of interpersonal respect. First, self-respect requires that we not degrade ourselves, or allow others to have undue control over our lives. But whereas interpersonal respect largely requires that we not interfere in the lives of others, this noninterference would be sorely out of place as a requirement for self-respect. Self-respect, in fact, often demands strong interference; to respect myself I ought to scrutinize my decisions, and hold myself to appropriately stringent behavioral standards.63 We can explain this asymmetry by considering what it takes to respect a person as a rational agent. It would be disrespectful of me to coerce you into making the decisions that I want you to make, as doing so would undermine your ability to exercise your rational agency. But there is no self-regarding

62 For more on symbolic acts of respect, see Joseph Raz, Value, Respect, and Attachment (Cambridge: Cambridge University Press, 2001), 172-175.

63 For more on this point, see Robin Dillon, ‘How to Lose Your Self-Respect,’ American Philosophical Quarterly, 29 (1992), 133-134.
parallel to this sort of paternalism: I might be able to force myself into doing something I do not want to do (say, by setting a noisy alarm to get myself out of bed earlier than I would like), but this self-interference is just another expression of the rational agency that I am meant to be respecting.

Section 2: Respect, Practical and Epistemic

Let us now take a closer look at the sort of rationality that lies at the center of the rationalist account of respect. No discussion of respect for persons would be complete without a mention of Kant, and so it is with him that we will start our investigation:

The sole object of respect is the [moral] law—that law which we impose on ourselves and yet recognize as necessary in itself...All respect for a person is actually only respect for the law (of righteousness, etc.,) that that person exemplifies.64

For Kant, it is the supreme law of our practical, rather than theoretical, reason that inspires within us a feeling of respect. And so, respecting persons is a matter of respecting their status as practically rational agents capable of willing the moral law.

In many regards, modern philosophers have moved away from the Kantian account of respect. While Kant spoke of it as a feeling or inclination, Darwall and others have emphasized the role that proper judgment and behavior play in respect. And yet, the Kantian conflation between rationality and practical rationality persists throughout the literature. According to Thomas E. Hill, respecting someone as an autonomous being (and thus, as a person), requires us “to acknowledge that certain decisions are up to him or her and thus to refrain from efforts to control those decisions.”65 Likewise, to respect oneself on Darwall’s

64 Kant, *Groundwork*, 4:402n.

account requires the recognition of oneself as “a person, a being with a will who acts for reasons.” Elizabeth Telfer characterizes disrespectful behavior as that which makes us look “less than human, in the sense that a human being is characteristically an autonomous being” and Joseph Raz understands respect as a special type of valuing that we extend towards valuers, where to be a valuer is to have the capacity to “one’s conduct and one’s life, in light of one’s understanding of those features of the world which are reasons, rather than merely responses to factors which are in fact reasons”. Even philosophers skeptical of Kant’s abstract and individualistic notion of autonomy are still loath to abandon the rationalist picture entirely. While Margaret Farley stresses that rationality “was never as pure and impartial as [Kant] claimed it to be,” she remains committed to the idea that autonomy (albeit, a type of autonomy that emphasizes the extent to which people are interdependent) is the feature of personhood that makes us deserving of respect.

And yet, practical rationality is not the only kind of rationality. In addition to being practically rational agents, we are also epistemically rational agents capable of weighing evidence, making inferences and drawing conclusions. As is the case with our practical rationality, our epistemic capacities are such that we can better or worse at exercising them. I can make bad decisions about how to spend my time, and I can also make bad inferences.

---


68 Raz, *Value, Respect, and Attachment*, 153.

69 See Margaret A. Farley, “A Feminist Version of Respect for Persons,” *Journal of Feminist Studies in Religion* 9 (1993): 198. Farley’s project in that paper is to retain the basic commitment to autonomy as the feature of personhood that makes us deserving of respect, while at the same time emphasizing the extent to which we are beings situated within a web of unique relationships and experiences.
towards unsupported epistemic conclusions. Here, we must ask: if a person’s status as a practically rational agent commands respect, why doesn’t their status as an epistemic rational agent?

To answer this question, a rationalist might point to an instrumental relationship that obtains between practical and epistemic rationality. Epistemic rationality is essential to practical deliberation; I could not make even the simplest of decisions without some understanding of how causation works. And so, since respecting you requires me not to impinge upon your ability to practically deliberate, it will by extension require me not to impinge upon your ability to make inferences and draw conclusions. This means that there is no point in talking about the respect-worthiness of a person’s epistemic rationality apart from the respect-worthiness of their practical rationality. Whatever reasons we have to respect them as epistemically rational agents will ultimately bottom out in the reasons that we have to respect them as practically rational agents.

The answers that Kant could give for the conflation are not available to the modern rationalist. First, Kant could give a phenomenological reason for denying that epistemic rationality ‘commands’ respect. Recall that Kant often treats respect as a non-empirical ‘feeling’ given to us a priori in our cognition of pure practical reason. *Groundwork*, 5:73, 5:76. Theoretical reason, simply as a matter of phenomenological fact, fails to inspire within us any similar non-empirical feelings. But this sort of explanation is lacking. Consider the feeling that one often gets when reading about a new discovery about the cosmos or solves a particularly challenging logic puzzle. The feeling of ‘exercising’ one’s epistemic capacities can come, I think, very close to the exhilarating awe that Kant sees as accompanying cognition of the Moral Law within.

But perhaps epistemic rationality simply does not need to ‘command’ respect. Respect, for Kant, was sometimes understood as an incentive that moves us to act in accordance with the demands of the Moral Law. And it makes sense why moral agents such as ourselves would need such an incentive. We are, after all, morally fallible agents capable of being moved by self-love as well as reason. And so if reason is to stand a fighting chance against desire, it will need to be able to pack a motivational punch. Theoretical reason, by contrast, requires no such incentive; there are no secret impulses of self-love lurking in Kant’s theoretical domain. We have, as Raz, *Value, Respect, and Attachment* 134–35 has noted, reason to be skeptical about this explanation. Self-love, wishful thinking, and vanity, after all, can and do pervert our theoretical judgments. Anyone who has ever had to begrudgingly acknowledge a flaw in their favorite argument can attest to how difficult it can be for theoretical reason to overcome self-love.
To see an example of this sort of ‘instrumentalist’ account, consider Hill’s classic explanation of the wrongness of benevolent lies in ‘Autonomy and Benevolent Lies.’

Some lies are clearly wrong because they’re vicious or self-serving. But not all lies have these traits. Instead, some lies are told out of a sincere concern for their hearer’s wellbeing. What is so wrong about these sorts of lies? Benevolent lies, Hill contends, are morally wrong because they “deprive people of a realistic picture of their situation” in a way that frustrates their ability to make fully informed life choices. The son who lies to his elderly, terminally ill mother about her prognosis, for instance, denies her the chance to make decisions about whether or not to pray, and about how to approach her impending demise.

71 Thomas E. Hill, “Autonomy and Benevolent Lies,” *Journal of Value Inquiry* 18 (1984): 251–67. ‘Instrumentalist’ accounts appear throughout the value theory literature. Julie Kirsch, in “What’s So Great About Reality?” *Canadian Journal of Philosophy* 35 (2005), 407–427, argues that self-deception is morally objectionable because it undermines our capacity for autonomous self-governance (p. 421). Barbara Herman, meanwhile, notes that effective deceit “causes the victim to have the beliefs necessary for her to adopt ends and choose actions that serve the deceiver’s purposes.” *The Practice of Moral Judgment*, (Cambridge: Harvard University Press, 1993), 228. Sarah Buss takes a different, albeit still instrumentalist, approach in “Valuing Autonomy and Respecting Persons,” *Ethics* 115 (2005): 195–235). While she denies that manipulation and coercion are wrong because they frustrate an agent’s capacity for autonomous choice, she nevertheless claims that they are wrong because they can negatively impact an agent’s welfare, prevent her from standing on equal ground with her manipulator and impede her ability to govern herself with an accurate understanding of her situation (226–27). Shiffrin’s recent *Speech Matters: On Lying, Morality, and the Law* (Princeton: Princeton University Press, 2014) offers a particularly nuanced instrumentalist explanation of why lying is wrong. Sincere communication enables us to “form and execute complex cooperative plans, to understand one another, to appreciate and negotiate around our differences, and to gauge, somewhat, the extent of our mutual ineffability” (1), which in turn help us fulfill our moral duties and grow as moral agents. When we lie, we frustrate humanity’s collective interest in maintaining reliable interpersonal channels of communication (23–24).

Miranda Fricker is a notable exception within this literature. In *Epistemic Injustice: Power and the Ethics of Knowing* (Oxford: Oxford University Press, 2007), she contends that we have both instrumental and non-instrumental reasons to see testimonial injustice (i.e. the undervaluing of an agent’s testimony as a result of prejudicial beliefs about her) as morally wrong. Testimonial injustice can cause its victims to lose confidence in their epistemic capacities and suffer psychological anguish, both of which frustrate their ability to get ahead in life. But testimonial injustice also wrongs agents in their capacity as knowers, which is “a capacity essential to human value” (44). Fricker does not, however, explain why this capacity gives rise to moral requirements, or whether respecting an agent as a “knower” involves more than according their testimony the value it is due.


73 Ibid, 262.
I find Hill’s picture both compelling and incomplete. It is compelling because it highlights the extent to which respecting an agent’s practical agency may require us to privilege that agency over their happiness or comfort. But, for reasons that I will make clear in this section, I think that Hill’s account overlooks a more basic reason that we have to be morally concerned about telling even the most benevolent of lies. Lying is wrong, on my account, because it frustrates our ability to exercise our epistemic rationality. And this frustration is wrong in and of itself, and not simply because we need our epistemic capacities to be in working order for practical deliberation. It is, in other words, morally wrong to frustrate an agent’s ability to understand the world around her in the same way that it is morally wrong to frustrate her ability to make decisions about her life. In both cases, the ‘wrongness’ can be located in the simple fact that she is a rational agent whose reasoning capacities deserve respect.

To see exactly where I depart from the ‘instrumentalist,’ consider what reasons someone would have to feel anger upon finding out that she has been told a benevolent lie. An example will help clarify the situation:

Hank discovers that his daughter, Gracie, has been rejected from her dream university. He discovers this before she does, and so he decides to lie to her and tell her that she was, in fact, accepted. Hank lies to Gracie with full knowledge that Gracie has already accepted a spot at a local college. He knows, in other words, that even if Gracie had been accepted at her dream school, she would have turned down the offer. Further, he lies with only the best of intentions—he hopes that his lie will give her a much-needed confidence boost heading into her freshman year.

Suppose that Gracie finds out about Hank’s lie several years down the road. There would be nothing unreasonable about Gracie if she were to react with anger towards Hank at this new discovery. But what justifies this anger? Hill is right to note that many of the ‘usual’ reasons that we have to be angry at people who lie to us do not apply in cases of benevolent lies. Gracie cannot be angry at Hank for acting maliciously or for causing her harm, because he
did neither of these things when he lied to her. But we also cannot appeal to Hill’s account to explain why Gracie would be justified in feeling angry at Hank. After all, Hank’s lie could not have impacted Gracie’s decision about what college to attend, as that decision had already been made at the time that he told the lie. And yet, I think that Gracie could still be justifiably angry at Hank for the simple reason that he made her believe something false about her life.

The instrumentalist could push back here: what matters is not that a lie will impinge upon an agent’s ability to exercise her practical reason. Rather, what is at issue is the fact that a lie could, however improbably, impinge upon an agent’s ability to make decisions for herself.\textsuperscript{74} Hank may end up being right in thinking that his lie will not impact Gracie’s decision-making process in the slightest, but there is no way that he can know this with absolute certainty. But we can tweak the case, albeit in a fantastical way, to get around this objection. Suppose that Hank is clairvoyant, and that Gracie has already put down a large, non-refundable deposit at her local college. Even if Hank knows that the rejection letter would never have impacted Gracie’s deliberation, and even if he also knows that telling her the truth about it would only cause her anguish, I still have the intuition that he should not hide it from her. His daughter deserves to know the truth about her university application results, even if this knowledge will not factor into her practical deliberations, and will only cause her distress.

This conclusion reflects what I take to be a deeply ingrained human commitment to knowing the truth for its own sake, and not simply for the role that it plays in our practical

\textsuperscript{74} This is in fact, precisely the move that Hill makes. Benevolent lies, Hill warns, “do not necessarily or always violate the right of autonomy, but we should not be hasty in concluding that a particular lie does not concern any significant decisions” (“Autonomy and Benevolent Lies,” 262).
deliberations. Consider the value that we often place on scientific discovery. Whether the earth is flat or round has no impact on the decisions I make about my life. I’d hope, however, that you would take the time to correct me if you found out that I mistakenly believed the earth to be flat. In the same spirit, there isn’t much difference between the moral decisions that I, a Kantian, actually make, and the moral decisions that my virtue ethicist friend makes. We both come out against police brutality and unfair voter ID laws. We both lie (sometimes), choose the expedient over the right (frequently) and stand up for what we believe in (not often enough). Perhaps we occasionally appeal to different standards when puzzling through a moral dilemma, but most of the time we make moral decisions by some combination of intuition, habit and convention. Nevertheless, I still care about having the right higher-order moral commitments. My desires to ‘be right’ about morality here do not reflect a deeper commitment to being a practically efficacious agent. Rather, it reflects a deep commitment to being able to know how things really stand.

We care about epistemic discovery in very much the same way that we care about practical deliberation. When I make choices for myself, I certainly want those choices to be good ones. But I want you to respect my choices not because they’re good ones, but because they’re mine. In fact, I hope that you would respect them even if you thought that they were bad choices. This is because, as Sinatra aptly put it, I attach non-instrumental value to being

---

75 This does not, of course, mean that you must wholeheartedly endorse every foolish decision I make. You can certainly argue with me about joining a cult or giving all of my money to that Nigerian prince. But it would be wrong of you to lock me in your basement to prevent me from doing either of these things. Indeed, the cases where paternalistic interference seem most conscionable are typically those that involve non-rational agents (infants, the severely mentally disabled), very mild restrictions on autonomy (seatbelt laws), or paternalistic actions that are undertaken in the service of preserving an agent’s future rational agency (involuntary psychiatric holds for the severely mentally ill).
able to do things “my way.” In much the same way, when it comes to the epistemic domains that matter most to us, we want not only to have justified true beliefs, but also for those beliefs to be a product of our own reasoning power. Most theists and atheists alike, for instance, want their theological commitments to be the product of introspection and deliberation, rather than cultural indoctrination.

**Epistemic Rationality and MoralDemandingness**

So far, I have highlighted one of the parallels between practical and epistemic rationality, *viz.* that we value both non-instrumentally. But I cannot jump from this result to the further conclusion that an agent’s epistemic rationality, like her practical rationality, gives me moral reason to respect her. One might worry here that making such a leap would have the effect of stretching the term ‘respect’ past meaningfulness. We can be said to value all sorts of different things for their own sake—art, the natural world and dead language-learning might all been considered non-instrumentally valuable. But I can think that there is something non-instrumentally valuable about learning Ancient Greek without committing myself to the idea that I am disrespecting myself if I never get around to learning it.

Does our epistemic rationality deserve respect in the way that our practically rationality does? Or is it more like learning Ancient Greek, in the sense that we can value it for its own sake while still denying that its value generates any moral requirements? To answer this question, let us first identify a salient difference between dead languages and practical rationality. One way to locate this difference is by first considering how we might motivate a rationalist account of respect for persons. We might do this is by employing a sort

---

of conceptual ‘opt-out’ test. One way to motivate the rationalist account of respect is by employing a sort of conceptual ‘opt-out’ test. If we’re trying to figure out whether some attribute is central to the concept of respect for persons, we can ask ourselves whether it would be possible to have respect for persons while denying the value of that attribute. Take practical rationality. Were a dictator to say “I respect the people, but I see no reason to allow them to make decisions for themselves,” you would be right to wonder whether he actually understood the meaning of ‘respect’ and ‘people’. The incoherence is even more obvious in the case of self-respect. Hill’s deferential housewife who, as the name implies, defers to her husband on practical matters, might be able to have all sorts of positive attitudes towards herself, but she cannot have self-respect. But ‘opting out’ of learning Ancient Greek gives rise to no similar incoherence. Rather, I can respect myself as a person without ever trying to read Homer in the original or memorizing the forms of an irregular Greek verb.

Does epistemic rationality look more like practical rationality or Ancient Greek in this regard? To answer this question, let us imagine what an epistemic parallel to Hill’s deferential housewife would look like. In Hill’s original example, the deferential housewife allows her husband to tell her what to do with her time, whom to befriend and what to wear. In the epistemic parallel, our deferential housewife—let us call her Allison—retains her ability to make decisions about how to spend her time and how to dress, but decides that she would be much better off to allow her husband to tell her what to believe. If Allison believes in evolution, it is only because he told her to that she ought to. On those rare occasions when

---

she draws her own epistemic conclusion on some matter, she always regards it as tentative: she will believe it, but only until her husband tells her otherwise.

Why is Allison so deferential? There are a few possible explanations. On the one hand, there is nothing inherently wrong with ‘giving up’ on drawing epistemic conclusions within a given domain because you rightly believe yourself to lack the ability to successfully reason within that domain. I am happy to defer to the scientists on matters about climate change, but I do not think that my epistemic deference denotes a lack of self-respect. Rather, my epistemic deference is much like my practical decision to let my accountant tell me how to do my taxes. Respecting myself as a practical agent, after all, does not require that I exercise my practical agency within every possible domain. For the same reason, respecting myself as an epistemic agent should not require me to spend hours reading biology textbooks in the hopes of someday being able to decipher the scientific literature on climate change. At most, it requires me to carefully scrutinize my decisions to opt out of certain epistemic domains.

If I can respectfully defer to the scientists on matters of climate change, why can’t Allison respectfully defer to her husband on epistemic matters more generally? Let us consider what it would take for Allison to carefully scrutinize her decision to indiscriminately opt out of epistemic domains. Suppose Allison decides to opt out of forming her own opinions after a great deal of personal reflection and deliberation. She has looked at the opinions that she has held in the past, and found them all to be the product of knee-jerk reactions and sloppy reasoning. After all this, she concludes that she is not good at this theoretical reasoning business, and that she is better off letting her husband tell her what to think.
But note that in performing the appropriate due diligence, Allison has inadvertently affirmed the very capacity within herself that she has intended to deny, *viz.* that she is an epistemically rational agent capable of critically assessing evidence and drawing reasonable conclusions based on that evidence. She may not have shown herself to be a particularly good epistemic agent—she could still be absolutely correct in her assessment of her previous epistemic judgments. Nevertheless, she has proven herself capable of independent thought. She has, in other words, demonstrated herself to be a proper object of epistemic self-respect.

Suppose Allison does not do her due diligence. Suppose her decision to opt out was made hastily and without much thought. Perhaps she opts out because she cannot be bothered to form her own opinions, or because she has internalized some deeply sexist beliefs about the rationality of women. Here is a situation where Allison’s decision to opt out is itself an act of disrespect. The capacity for independent thought is a valuable human ability, and Allison has tossed it away carelessly or out of a deeply internalized, but also deeply mistaken, sense of her epistemic worth. Were we to meet Allison, we might be tempted to implore her to have some self-respect and stop letting her husband do her thinking for her.

But there is another way that one might opt out of epistemic rationality. In both of the above scenarios, Allison still cares about epistemic rationality in some sense. She wants someone to be in the business of forming good opinions on her behalf, but does not want it to be her. What if Allison simply denied that epistemic reasoning and the pursuit of truth have the value that I purport them to have? Could she then opt out of caring about epistemic rationality while still respecting herself? Consider a variation on the original case:

Allison has left her husband and joined a clique of radical epistemic nihilists. There is no point, Allison thinks, in learning about the world or trying to arrive at true beliefs, because the world is what you make of it and truth is an illusion.
What would take for Allison to perform ‘due diligence’ in this scenario? In one version of the story, she formulates arguments against an objective external world and comes up with evidence in support of the claim that truth is a ‘matter of opinion’. Here, Allison will have, by due diligence, inadvertently affirmed that she holds herself to the very sorts of objective epistemic standards that she, as an epistemic nihilist, claims not to exist. So long as she makes that affirmation, and so long as she continues to see herself as an agent capable of reasoning in accordance with those standards, Allison will be unable to coherently and respectfully opt out of valuing her epistemic rationality.

But imagine that Allison’s epistemic nihilism comes from some source other than philosophical argumentation. Perhaps she is depressed after her divorce, or perhaps she enjoys being a contrarian. Allison can choose not to respect herself as an epistemic agent—but so long as she remains epistemically rational, her choice will constitute a failure of self-respect.

To try to respect oneself without respecting one’s ability to make epistemic judgments is either hypocritical or self-defeating. Thus, just as respecting a person requires me to respect them as a practical agent, so too does it require me to respect them as an epistemic agent. And, just practical rationality is worth respecting for its own sake, so too is epistemic rationality worth respecting for its own sake.

Section 3: The Requirements of Epistemic Respect

My ultimate aims in this chapter are to develop an account of epistemic self-respect, and to consider what relation, if any, obtains between epistemic self-respect and self-knowledge. Before I do this, however, I will first say something about what it would look like to treat another person with epistemic respect. For the rationalist, “practical rationality”
has historically been the answer to both the question “why should I respect a person \textit{qua} person?” and the question “what is it that determines whether an action is respectful or disrespectful?” But practical rationality is not the only sort of rationality that gives us reasons to respect persons. This raises a question: if I owe epistemic respect to myself and others, what exactly does this obligation look like?

First, a word about recognition respect more generally. Recall that, for Darwall, recognition respect requires both proper judgment and proper self-constraint.\footnote{Darwall, \textit{Respect}, 36-49. Darwall thought that we could also have recognition respect for laws, institutions and people by virtue of the roles that they occupy. If I ought to have recognition respect for Beverley McLachlin because she is a Canadian Supreme Court Justice, then I should also have recognition respect for Russell Brown.} Darwall discusses these requirements as they relate to practical rationality, but I believe that they are equally applicable to the case of interpersonal epistemic respect. As with practical respect, epistemic respect requires not only that I be willing to factor your status as an epistemically rational agents into my deliberations, and restrict my actions towards you accordingly, but also that I do so for the right reasons. I should judge you to be deserving of epistemic respect simply by virtue of your status as a person, and not by virtue of your membership in a certain race, sex or social class.

As in the case of practical respect, the importance of proper judgment can be made clear with an example. Suppose that Alistair is an incorrigibly sexist physics professor who ardently believes that women lack the capacity to understand high-level physics. Now, suppose that one day, Alistair makes a bet with his colleague about a female undergraduate named Susan. His colleague bets Alistair that, if he were to hold Susan to appropriately stringent epistemic standards, he would find her to be as capable of understanding the
material as his male students. Alistair takes the bet and proceeds to treat her like ‘one of the boys’. Susan may not be able to complain about how Alistair has treated her as part of the bet, but, were she ever to find out about Alistair’s opinion about female rationality, she could justifiably object that he has failed to respect her as an epistemically rational agent.

Second, epistemic respect requires that we behave in a way that assumes persons to be, absent extenuating circumstances, capable of ‘handling,’ and invested in knowing the truth about their lives and the world around them. At a minimum, then, we ought to view lying and deception as things that ought, generally, to be avoided. This is true even when it comes to cases of benevolent lies and well-intentioned deceptions. I may tell my younger sister that our parents really did take Fido to a farm upstate to save her some grief, but in doing so, I fail to respect her an epistemic agent. When we protect people from hurtful truths, we place something—that is, their feelings—above their status as epistemic reasoners. This sort of prioritizing may be consistent with caring deeply about someone, but it is rarely consistent with respecting them.

Epistemic respect will thus sometimes require us to hold people to sufficiently stringent epistemic standards. That respecting persons involves more than self-constraint a point that Darwall’s original account of recognition respect misses. Nevertheless, in both the practical and epistemic realms, respecting persons will sometimes require us to be willing

---

79 As Fricker, Epistemic Injustice, 86–108, notes, this will often require us to reflexively assess our own epistemic shortcomings. We occupy various social roles that can have a distorting effect on our perception of other agents’ epistemic capacities, and epistemic justice, on her account, demands that we take stock of, and make efforts to correct for, these distortions.

80 This ‘positive’ requirement of respect could be re-translated into a constraint: one ought to constrain oneself from holding persons accountable to inappropriately low standards. Nevertheless, it seems more natural to speak of respect for persons as requiring us to both act and refrain from acting in certain ways towards them.
to make suitably stringent demands of them, and to hold them accountable for meeting or
failing to meet those demands. To see this feature of respect as it relates to practical agency,
consider the following case:

Bill has hired Cynthia to work as the receptionist at his law office. He has a generally
low opinion of women. While there might be some exceptions, Bill thinks that, as a
rule, women aren’t capable of very much. As such, he normally makes an effort to
hire male receptionists. This time around, all the candidates were women, and
Cynthia seemed, in Bill’s eyes, to be the best of the bad bunch. As it turns out,
Cynthia really is awful at her job. She isn’t incompetent—she’d be a perfectly fine
receptionist if she just tried. But Cynthia doesn’t care to be a fine receptionist, and
she lets this fact be known. She is rude to clients, she misfiles important paperwork,
she takes suspiciously long lunch breaks and she shows up to Bill’s very formal office
in ripped t-shirts and party dresses. If Cynthia were a man, Bill would have fired her.
But Cynthia is not a man, and in Bill’s eyes, she’s doing the best a woman can be
expected to do. So he keeps her employed, and gets his legal assistant to pick up some
of Cynthia’s slack.

Bill has disrespected Cynthia twice over, and I don’t think either instance of disrespect is a
matter of a failure to properly constrain his behavior. First, his views about women are
themselves disrespectful to Cynthia. She’s a person, and he ought to recognize that her status
as a person makes her the proper object of recognition respect. Second, Bill has held her to
standards that are inappropriately low for the sort of entity that she is. Cynthia is a rational
agent in control of behavior, and this means that she ought to be held accountable for the
decisions that she makes. This second form of disrespect is, in this case, an offshoot of the
first: if Bill respected Cynthia as a person, he would be willing to hold her to accountable for
her poor performance. Nevertheless, the two forms of disrespect can come apart from each
other. Were Bill to hold Cynthia to appropriately stringent standards simply because he
wanted to prove that women weren’t capable of being good employees, he would still be
guilty of the first type of disrespect. Likewise, had Bill’s lenience been born out of a personal
aversion to hurting others, rather than a misogynistic worldview, his failure to hold Cynthia
to appropriately high standards would still have been disrespectful. For this reason, I think it is best to amend the definition of recognition respect once again.

And so to when it comes to respecting people as epistemic agents are we often required to hold them accountable to suitably stringent epistemic standards. To truly respect his students, for instance, a professor must be willing to penalize them for sub-par work. To do otherwise—to indiscriminately assign high grades to one’s students out of the belief that they couldn’t possibly produce work actually deserving of those grades, coupled with the desire to be a nice professor—demonstrates an unwillingness to take them seriously as epistemic agents capable of exercising that agency in better and worse ways. Likewise, when a doctor tells her patient that he ought to follow her instructions “because she said so” instead of explaining to him how her treatment plan is meant to work, she disrespectfully assumes him to be incapable of understanding even basic medical reasoning. Her assumption might turn out to be correct, but given that something as important as respect is on the line, the doctor has a good moral reason not to make it.

It is for this reason that respecting people as epistemic agents often requires us to take seriously their intellectual claims and opinions. To tune out your friend when she tells you all about how vaccinations cause autism is to write her off as a person whose opinions and beliefs do not matter. To try to intellectually engage with her, in contrast, is to hold her to

---

81 Here, one can’t help but think of the modern-day ‘helicopter parents’ who ardently protect their children from experiencing personal failure. These are the parents that complain to professors and deans whenever their college-age children receive bad grades or have disputes with their roommates. By offering this sort of ‘protection’, helicopter parents treat their children as though they are incapable of settling their own disputes or taking responsibility for their own shortcomings. In this way, helicopter parents hold their children to inappropriately low, and thus disrespectful, standards.
epistemic standards appropriate to her status as a rational agent, and to affirm that it matters that her opinions be reflective of that status.

This is not to say that respecting others requires us to constantly scrutinize their beliefs and opinions. Conversational norms are complicated, and sometimes when we have discussions, we do so with the presumption that our interlocutors are not holding their claims up for debate.82 When my aunt tells me at her husband’s funeral that “he’s with Jesus now,” she is not inviting me into a theological argument. In fact, respecting people may often require that we respect their opinions simply because they hold them by biting our tongues and nodding politely. By holding our tongues, we affirm that people are capable of, and deserve the opportunity, to form their own beliefs about the world around them.83

Epistemic respect, like its practical analogue, will thus often involve a precarious balancing act between active engagement and polite restraint. When we strive to be honest to others, and to intellectually engage with the claims that they make, we affirm that they are the sorts of being who deserve to have rationally defensible opinions and true beliefs, and who are epistemically sensitive to argument and evidence. When we appropriately restrain ourselves from this sort of engagement, in turn, we acknowledge that people deserve the

82 Shiffrin, *Speech Matters*, 16 calls these situations ‘suspended contexts’. We are justified in temporarily suspending our regular presumption of truthfulness in discourse, on her view, when doing so would “serve other valuable purposes whose achievement depends upon the presumption’s suspension” and when “the fact and justification of suspension is publicly available.” For example, the question ‘How are you doing?’ is not one that typically is asked with the expectation of a truthful answer. Instead, it is a mere formality that people use to acknowledge one another (33).

83 Jason N. Joh makes a similar point in “Respect for Persons,” *Theory into Practice* 14 (1975): 271–78. To respect a person requires that we recognize their opinions to be “personal, coming from him or her as a distinct center of consciousness, and taking them as results of distinct ways of experiencing the world and as such to be important” (273). Sometimes this will require us to critically engage with those opinions, and sometimes it will require us to refrain from engagement.
opportunity to make up their own minds about what to believe, even when they do so in a way that departs from our own understanding of the world.

**Section 4: The Requirements of Epistemic Self-Respect**

What does it mean to respect oneself as an epistemically rational agent? To answer this, let us take a closer look at the requirements of recognition self-respect more generally. First, recognition self-respect requires one to have made the right judgments about why one is deserving of respect. If I only see myself as worthy of respect because I am popular or wealthy, then I have not simply located an unstable attribute upon which to hang my sense of worth. Rather, there is an important sense in which I have failed to have respect for myself.

What else does recognition self-respect require? To see oneself as deserving of respect is, first, to have a certain self-understanding. The self-respecting agent knows that she deserves not to be treated in ways that are commonly considered dehumanizing. She also recognizes that she is entitled to make moral demands of others.\(^{84}\) Further, self-respect requires us to avoid treating ourselves in certain ways. Interpersonal respect may require me to respect your decision to spend your life engaged in a project that I consider trivial, but intrapersonal respect carries with it the exact opposite demand. The decisions that I make are the products of my rational agency and, as such, self-respect requires me to be invested in ensuring that they properly reflect the value of that agency. In both the practical and epistemic domain, there are standards that our decisions must meet in order to be considered

---

\(^{84}\) It is because of this sense of entitlement that self-respect has often been associated with reactive emotions like anger and resentment. I can only be resentful about your paternalistic treatment of me if I have a sense that my autonomy deserves respect. By the same token, I will only feel angry at you when you refuse to even entertain what I have to say on a subject if I have a sense that people ought to respect me enough to at least consider my opinions. For more on this point, see Dillon, “Self-Respect: Moral, Emotional, Political,” *Ethics* 107 (1997): 230, and Macalester Bell, “A Woman’s Scorn: Toward a Feminist Defense of Contempt as a Moral Emotion,” *Hypatia* 20 (2000): 84–85.
rationally defensible.\textsuperscript{85} Self-respect, I believe, requires a commitment to making choices and forming beliefs in a way that accords with these standards.

In one sense, this requirement of accountability is not unique to recognition self-respect. As I noted in the previous section, respecting other people can involve holding them accountable to sufficiently stringent epistemic standards. But our reasons for doing this in the two cases are subtly, but importantly, different. In the interpersonal case, I need not be invested in whether or not you actually \textit{meet} these standards. Rather, I need to hold you accountable for your beliefs only insofar as failing to do so would be tantamount to denying your status as an epistemic agent capable of better and worse forms of reasoning.\textsuperscript{86} The intrapersonal case is different than this. I have control over my beliefs and actions, and I bear a relationship to them that I do not bear to other people’s beliefs and actions. I am, as I will elaborate on in the next chapter, simply too close to myself to remain neutral on the subject. And so, returning to Darwall’s original distinction between recognition respect and appraisal respect, we might say that while these two types of respect remain distinct in the interpersonal case, the boundaries between them become blurry in the intrapersonal case. To have recognition respect for myself as a person requires that I strive be the type of person for whom I could reasonably have appraisal respect.\textsuperscript{87} To do less than this, and to live my life in

\textsuperscript{85} I do not aim to give an account of what these standards are. Nevertheless, a few commonsense candidates immediately come to mind: I shouldn’t believe \( p \) and \( \neg p \), and if I discover that I do believe \( p \) and \( \neg p \), I ought to endeavor to determine which proposition I ought to reject; I should revise my beliefs in light of new, relevant evidence, I shouldn’t believe \( p \) simply because I desire \( p \) to be true, and so forth.

\textsuperscript{86} There are exceptions here. If I am holding my child or friend accountable to these standards, I can be doing so not simply because I ought to, but also because I am invested in helping them improve as rational agents.

\textsuperscript{87} I owe this point to Michael Cholbi.
a way that is undeserving of positive appraisal, is to show a disrespectful indifference to the very rational agency that makes me deserving of recognition respect in the first place.

This is not to say that self-respect requires that we care equally about every choice we make and opinion we form. Just as we don’t lose our self-respect every time we make a subpar practical decision, so too is epistemic self-respect consistent with a certain amount of epistemic negligence. I can’t possibly scrutinize every belief that I have, and often considerations about expediency will move me to form beliefs based on evidence that I recognize to be somewhat insufficient. Further, I can’t be expected to care equally about every one of my beliefs: to explain why I am not at all motivated to double-check whether I’m right in my belief that Alan Jackson is Canadian, I need only point to my indifference to country music, and to the complete irrelevance of Alan Jackson-related facts to my life.88 Rather, as Dillon describes it, for an agent to have respect for herself as a person, she must endeavor to live a life governed by the “the ideals, aspirations, commitments and “points of no return” that define her as the particular person she is.”89 Self-respect involves a sensitivity to one’s status not simply as a generic rational agent, but also to one’s status as a unique individual with specific commitments, abilities and ends. These commitments, abilities and ends are, in large part, manifestations of my rational agency; they are the result of choices that I have made or failed to make, propensities that I have cultivated or failed to cultivate and so forth. Given this, I think that self-respect requires that we be appropriately invested in the ends that we, as particular individuals, set for ourselves.

88 Of course, after writing this sentence, I felt the urge to double-check Alan Jackson’s birthplace. As I discovered, he is actually American.

Before I discuss this last requirement in greater detail, I must say something about what I mean by an ‘end.’ Typically, when philosophers talk about ends, they have in mind practical ends. I can choose to play a sport, cultivate a new friendship, or plan a trip to France. But we can also set epistemic ends for ourselves. We can choose to direct our attention towards certain epistemic domains and away from others. Often, our reasons for choosing one subject of epistemic inquiry over another is pragmatic. If I want to get into law school, I better learn how to do those pesky LSAT logic games. In other cases, our practical aims can only be fully explained with reference to our epistemic interests. There would be something sad about the person who dedicated his twenties to learning philosophy simply because he desperately wanted all of the lifestyle perks that come with being a philosopher professor. Instead, one would hope that the practical end that he has set for himself (‘becoming a philosophy professor’) is to be explained by his epistemic interest in doing philosophy. Although we can set all sorts of ends for ourselves, I will be especially concerned in this discussion with the ends that I have called ‘identity-constituting ends’. These are the ends that speak to our sense of who we are as individuals; they often relate to our values, social identities, personal projects, and most meaningful relationships. Why does self-respect require that I invest myself in realizing my ends? Consider what it means to fail to be invested in one’s ends. We might imagine someone who is always willing to give up her personal projects when they conflict with those of her husband. She may plan to join a book club, but if the club meetings are on the same night as her husband’s intramural hockey games, she will put down *Atonement* and head to the rink. By constantly giving up on her own projects, the woman expresses something about how she sees them (as less important than her husband’s), and also about how she sees herself (as less important than her
husband). But this sort of self-image is antithetical to recognition self-respect; to respect herself as a rational agent requires that she see herself as having the same worth as every other rational agent. Just as I would be disrespecting someone if I forced them to give up their own personal projects simply because they conflicted with my own, so too would I be disrespecting myself if I perpetually subordinated my own ends to those of another. And likewise, just as self-respect often requires that we demand of others that they take our ends seriously, so too does it require that we demand of ourselves that we take those ends seriously.

As the above example illustrates, proper self-investment will require that we not give up our ends too quickly. They are, after all, our ends, and this fact gives us the same reasons to respect them as it does everybody else. But investing in one’s ends involves more than just a willingness to pursue them. Specifically, self-respect will often require that we be willing to hold ourselves to appropriately stringent standards with regard to our ends. Massey captures this point vividly when he considers the case of a professor deliberating about whether to appear as a guest on a television quiz show:

Consider Professor Andersen, who has spent her life studying and teaching Elizabethan literature, and is proud of her scholarship and the love of literature she has engendered in her students. The producers of “Whiz-Bang” quiz show, a midday extravaganza, make Professor Andersen a lucrative offer to appear as a guest, for they think that she will help to establish that their show is superior to its competitors. Out of respect for herself (and her profession), Professor Andersen refuses the offer. She is not averse to money. Rather, she believes that the use of a reputation based on the pursuit and dissemination of learning to legitimate a television quiz show could demean herself and prostitute her values.90

Whether or not profiting from a quiz show really does constitute a prostitution of one’s professorial powers is up for the debate. But what is not up for debate is the fact that there are better and worse ways to pursue one’s individual ends. Since respecting myself requires proper investment in my ends, I have moral reason to care that I be pursuing those ends in accordance with sufficiently stringent standards of behavior and reason.

In order to demonstrate how epistemic self-respect can require us to be invested in holding ourselves to stringent standards when it comes to our epistemic ends, I will consider at length what it means to respect oneself in the context of one’s philosophical ends. I will suggest that, insofar as epistemic rationality is itself worthy of respect, and insofar as philosophy is an epistemic end that we can set for ourselves, then proper philosophical engagement can be a requirement of self-respect.

*Two Philosophers*

Consider the following two fictional, but perhaps familiar, philosophers. Felix is senior philosopher at a ‘Leiteriffic’ department. He has published dozens of highly-regarded papers and has earned a reputation as an inventive and rigorous thinker. Unfortunately, in the last five years, Felix’s reputation has gone to his head, and his work has suffered accordingly. Simply put, Felix has become so sure of his philosophical prowess that he no longer pays attention to the quality of his work. He publishes too quickly and too often, churning out papers that contain fallacious arguments and under-explored philosophical commitments.

---

91 Massey (ibid, 248–49) himself takes those standards as proof that the requirements of self-respect cannot be wholly objective. I do not think this is a bullet that ‘objectivists’ about self-respect have to bite. One can still say that there are objective standards that a behavior must meet in order for it to qualify as a requirement of a self-respect, while allowing that individuals have a great deal of latitude as to what sorts of behaviors they come to see as requirements of self-respect.
Now consider Felix’s colleague, Sarah. Sarah is an early-career philosopher who, like Felix, enjoys a reputation as an inventive and rigorous thinker. She has only published a few papers, but she is already well-regarded by her peers. Unlike Felix, however, Sarah suffers from crippling self-doubt about her philosophical abilities. She sees herself as a charlatan, entirely undeserving of the professional success that she has enjoyed. This has a tangible impact on her work. Sarah frequently finds herself abandoning philosophical commitments when she discovers that they are not shared by others in her field, and when someone raises an objection to one of her arguments during a conference talk, Sarah simply admits defeat.

There is something unfortunate about both Felix and Sarah, but what is it? In this chapter, I argue that Felix and Sarah are both failing themselves in a similar way, and that these failures are distinctly moral failures. Specifically, both philosophers are failing to respect themselves as epistemic agents.

These cases were chosen quite intentionally. First, they should strike the reader as familiar. Even if we have not personally been Felix or Sarah, we all know people like them. Second, these cases are both ‘non-moralized.’ By this, I mean that no great harms or rights-violations are going to result from these agents’ epistemic failures. The worst that will likely happen as a result of Felix’s philosophical lackadaisicalness, for instance, is that some other philosophers will be stuck reading subpar articles. Presumably, whatever grievance one can justifiably have towards the author of a subpar philosophy article does not rise to the level of a rights violation. Likewise, while it is certainly unpleasant to admit to the failings of one’s argument mid-way through a conference Q&A, it is hard to see how doing so could be considered morally wrong. I hope to show that these cases both involve agents who have
failed to respect themselves epistemically. To begin this task, I will take a closer look at each case.

**Felix**

Let us start with Felix. I think most philosophers would agree that there is something problematic about being indifferent to the quality of one’s philosophical work. And indeed, there are many different places where one might locate this wrong. Felix’s philosophical peers could legitimately complain about his philosophical carelessness. He has, after all, added to the glut of bad philosophy papers that other people have to read.

But these considerations do not exhaust what is wrong with Felix. The problem with Felix, after all, is not simply that he has produced bad work. To say that a philosopher disrespects his epistemic rationality whenever he produces a sub-par argument would be to hold him to an epistemic standard that is far too stringent. Even good philosophers make bad arguments once in a while.

Or we might be tempted to say that Felix has disrespected philosophy itself. Philosophy aspires to truth and rigor, and Felix’s work falls short of both of these aspirations. But this, once again, does not quite capture the case. The freshman student who hands in a poorly reasoned, sloppily written paper purely for the purposes of passing the course and fulfilling his philosophy distribution requirement has also failed to respect the philosophical pursuit of truth. But while the freshman’s paper, like Felix’s, lack philosophical rigor, I think our reactions in the two cases will be quite different. The freshman’s lack of philosophical engagement is regrettable, but understandable; we have all had that university course that just did not capture our attention. Felix, in contrast, is a professional philosopher. He has chosen to dedicate himself to philosophy, and he has proven himself to be philosophically capable in
the past. The main problem with Felix, then, is not that he is disrespecting philosophy, but that he is disrespecting himself as a philosopher.

What does it mean to respect oneself as a philosopher? For starters, it requires an investment in philosophy as a worthwhile epistemic project. But on top of this, respecting oneself as a philosopher requires one to care about how well their own work is contributing to that project. If a philosopher really wants to produce an argument that’s worth accepting, for instance, she should be willing to subject it to the same level of scrutiny that she would an opponent’s argument, and she should be open to abandoning it if it proves to be indefensible. She should, in other words, commit herself to producing arguments that are, by her own lights, worth endorsing.

Felix’s philosophical carelessness is problematic, then, not simply because it conveys a lack of respect for philosophy as a whole, but also because it conveys a lack of respect for himself as a philosopher. He is, after all, philosophically talented. He has produced good work in the past, and he ought to care about whether his recent work is meeting those same high standards. But Felix’s ego gets in the way of this. And in the process of continuing to feed it, he has stopped taking his work as seriously as he ought to.

Of course, there may be situations wherein setting philosophically low standards for oneself is consistent with respecting oneself as a philosopher. Consider the case of Cosima. Like Felix, Cosima routinely sends in sloppy philosophical articles for publication. But unlike Felix, Cosima is still on the tenure clock. Her department has informed her that they love her work and want to keep her permanently, but that a university-wide policy requires her to have a few more articles before she can be granted tenure. So she makes a deal with
herself: she will temporarily allow her scholastic standards to slip in order to meet her university’s requirement.

Cosima has set epistemic standards for herself that are not properly reflective of her status as a professional philosopher. But unlike in Felix’s case, I do not have the intuition that Cosima is disrespecting herself by adopting these standards. She is, after all, making a decision that will greatly impact her life. If she fails to get tenure at her institution, she may very well find herself permanently out of the philosophical profession. Her decision to temporarily set a lower epistemic standard for herself was not made lightly or in the service of some trivial end, but rather, was made in the service of accomplishing one of her life’s ambitions. To expect Cosima to maintain suitably high epistemic standards at the expense of one of her life goals is, I think, to unreasonably privilege her status as an epistemically rational agent capable of producing good philosophical arguments over her status as a practically rational agent capable of making decisions about how she wants her life to go.

Note that I can say all this about Cosima’s decision while maintaining that epistemic rationality is, in some sense, worth having for its own sake. To say that epistemic rationality is worth respecting for its own sake is not to imply that it has an endless value that can never be sacrificed in the pursuit of other values. Rather, to say that epistemic rationality is worth respecting for its own sake is to say that we have reasons to respect it that are independent of the reasons that we have to respect our practical rationality. Cosima’s epistemic rationality is something that is worth respecting for its own sake, but it is not something that is worth privileging above everything else regardless of cost.
Sarah

What about Sarah? Unlike Felix, Sarah actually holds her work to very high standards. She is cautious about the philosophical conclusions that she draws and she is open (in fact, too open) to criticism. She is, in other words, thoroughly invested in producing good work. And yet, I still think that Sarah lacks respect for herself. If I were to watch her give up on her own argument during a Q&A period, I might find myself inwardly cringing and wishing that she was a bit more self-confident.

Sarah is, in this way, the epistemic equivalent of the wife who is too willing to give up on her life projects in the service of her husband. The wife in this example might have very good personal aims—she might, in other words, be adept at setting good ends for herself—but her inability to invest in those ends reveals a lack of self-respect. She simply does not take her own ends, and by extension, herself, as seriously as she ought to.

To see this lack of seriousness in Sarah’s case, consider what it would take to respect another philosopher as an epistemic agent. To respect another philosopher, I must extend to her a measure of ‘epistemic good faith’—when this other philosopher makes an argument, for instance, it would behoove me to take that argument seriously and make a good faith effort to understand why she endorses it.\footnote{Cynthia Townley makes a similar point in “Toward a Revaluation of Ignorance,” Hypatia 21 (2003): 37–55. There, she argues that epistemic responsibility requires us to see others as epistemic agents with unique perspectives that we ought to strive to appreciate (44). While she acknowledges that the requirements of epistemic responsibility might intersect with our moral and political requirements, she ultimately sees them as distinct (37; 45). On my view, our epistemic requirements of respect are themselves moral requirements insofar as they concern the proper acknowledgment of a person’s status as a rational agent.} If Sarah respects her colleagues, then she will extend just this sort of epistemic good faith to them. And indeed, her behavior during her conference talks suggests just this. By conceding the points that her interlocutors make,
Sarah is acknowledging that she sees them as philosophers capable of formulating incisive objections. But by ‘giving up’ on her epistemic conclusions too easily, Sarah fails to treat herself with the very sort of epistemic respect that she willingly extends to her peers.

It is probably significant that Sarah is a woman in this example. As Borgwald has noted, systemic inequality can make it difficult or impossible for members of underrepresented groups to take their own ideas seriously. Borgwald describes the sorts of subtle social ‘corrections’ that can contribute to this systemic self-discounting:

As girls they are taught cognitive deference, trained to discount their own judgments and rely on those of their parents/elders (Nelson 1996). A young woman may be told “you don’t want to be a doctor dear, you want to be a nurse.” This is a way of denying that she has the desires she thinks she has, and so makes her doubt herself.93 A lifetime of these sorts of ‘corrections,’ Borgwald contends, can frustrate a woman’s ability to “think autonomously, reflect on and evaluate [her] emotions, beliefs and desires, and to trust those judgments rather than deferring to others.”94 For Borgwald, this loss of epistemic personhood is consequentially bad. To lack epistemic personhood is also to lack the ability to care about one’s wellbeing. But on my view, a loss of epistemic personhood is bad in itself precisely because it constitutes a failure to respect a feature of one’s personhood that ought to be respected.

**Conclusion: Epistemic Self-Respect and Self-Knowledge**

At this point, I have shown that our reasons for respecting ourselves as epistemic agents are of the same kind as the reasons we have for respecting ourselves as practical

---


94 Ibid, 73.
agents. In this way, epistemic respect is something that we can owe both ourselves and others. I have also demonstrated that, among the requirements of epistemic self-respect, is a requirement to take one’s epistemic ends seriously by holding oneself accountable to appropriate epistemic standards.

Let us now return to Didion to consider the relationship between self-respect and self-knowledge. For Didion, it was the discovery that she was not as exceptional as she believed herself to be that constituted the beginnings of genuine self-respect. How should we explain this transformation? Of course, we can grant that the very fact that Didion is willing to entertain beliefs about herself that she finds uncomfortable and undesirable evinces the respect she has for herself as an epistemic agent. Had she clung to her misguided sense of superiority even in the face of reasonable evidence to the contrary, we could rightly say that she was going too easy on herself as a thinker, and privileging her own comfort over the truth.

More fundamentally, we can note that recognition self-respect, in both its practical and epistemic form, requires that we act in the right sort of ways for the right sort of reasons. It is not enough for me to simply act respectfully towards you; rather, I ought to do so because I understand you to be deserving of respect as a person. While Didion may have treated herself with respect before the revelation, she did so out of a sense of her own exceptionality, rather than an appreciation of her personhood. What her self-revelation gave

---

95 I have assumed throughout this chapter that epistemic rationality and practical rationality are distinct forms of rationality. This may not be the case. Rather, it may be the case that there is only one kind of rationality which we can direct towards both practical and epistemic ends. If this is right, then the problem with the traditional rationalist picture of respect is not that it focuses too narrowly on the respect-worthiness of practical rationality, but that it focuses too narrowly on the respect-worthiness of persons as practical end-setters.
her was the ability to see that her respect-worthiness did not disappear along with her sense of inherent superiority. Rather, she discovered that she still had a reason to respect herself, not because she was an exception to the rules that applied to other people, but precisely because she was not.

We can say, then, that self-respect is incompatible with certain forms of self-ignorance. If I treat myself with respect because I think that my wealth, beauty, or smarts makes me worthy of it, then I have performed the right actions for the wrong reasons, and thus have failed to have recognition respect for myself. Rather, to have recognition self-respect requires that I have some appreciation of the fact that I deserve respect simply insofar as I am a person.

How contentful this belief needs to be is up for debate. Do I have to grasp specifically that my rational agency is the quality of personhood that makes me deserving of recognition respect? Must I have ever explicitly thought the thought that my personhood makes me deserving of respect, or is it sufficient that I act respectfully towards myself out of some vague sense of my inherent worth or the fundamental equality of human beings? And how explicit and clear must my sense of the value of my personhood be in order for it to count as a piece of knowledge? The answers to these questions are not obvious. I imagine that, like Didion, we often only explicitly question what makes us deserving of self-respect when we are struck by the realization that we aren’t as unique and wonderful as we previously believed. Nevertheless, however we settle this question, it is clear that it will not lead us to a positive duty to know ourselves as individuals. While Didion’s example provides evidence for the fact that there can be a positive value to self-knowledge even when it is emotionally upsetting, this value is obviously instrumental in nature. Insofar as upsetting knowledge of
our shortcomings can give us an accurate self-perception, and insofar as an accurate self-
perception can rid us of our false sense of superiority, which in turn can make us aware of the
proper grounds of our self-respect, then such self-knowledge is valuable.

Further, if we grant that self-respect constitutively requires one to have a certain
knowledge of one’s personhood, we still haven’t located a non-instrumental reason to seek
any knowledge about who we are as individuals. The knowledge that self-respect may
constitutively require is not knowledge pertaining to us as individuals, but rather knowledge
pertaining to us as generic members of the human species. We cannot appeal to the non-
instrumental value of this generic self-knowledge to explain why we ought to set knowing
our individual selves as an end.

What this account of self-knowledge can do is shed light on how we ought to pursue
the end of knowing ourselves, once we have set it. Specifically, it can make sense of why
there is a moral difference between the person who seeks knowledge of herself ‘warts and
all’ and the person who seeks self-knowledge only selectively, and with an eye to
maintaining a positive self-conception. As I will show in the next chapter, while it may be the
case that both of these people value substantial self-knowledge for its own sake, only one of
them holds herself accountable to appropriately stringent epistemic standards. The difference
between the two people, in other words, is one of epistemic self-respect: while the former
agent clearly has it, the latter does not.
CHAPTER 2: SELF-KNOWLEDGE AND THE REASONS OF LOVE

Introduction

Self-respect can provide us with some normative guidance as to how we ought to pursue the project of knowing ourselves, but it cannot tell us *that* we ought to know ourselves. If we are to secure that second ‘ought’, we must do so by appealing to something other than the requirements of self-respect.

This chapter, in a nutshell, aims to secure that second ‘ought’ by investigating the reasons that love gives us to know its objects. My argument will proceed as follows. After offering some clarifications about the project (Section 1), I will in Section 2 turn my attention away from self-knowledge and towards interpersonal knowledge. In doing so, I will consider the reasons that we have to know other people, and conclude that there is a class of relationships—which I uncreatively call ‘loving relationships’—that necessarily require a commitment to knowing the person with whom you share it for the sake of knowing them. This commitment is normatively binding: so long as we share a loving relationship with each other, I ought to be committed to knowing you for the sake of knowing you. I will then show in Sections 3 and 4 that there is a strong analogy between the reasons we have to know other people and the reasons that we have to know ourselves. We are in an inescapable relationship with ourselves that is answerable to norms of both self-love and self-respect. Self-love generates non-instrumental reasons to know ourselves, while self-respect demands that we take seriously those reasons by pursuing our project of self-inquiry in accordance with
suitably stringent epistemic standards. We are therefore morally required not just to think about who we are, but to think about who we are clearly and carefully enough so that we come to know ourselves.

Section 1: Clarificatory Remarks

Before proceeding to my positive project, I will offer two clarificatory remarks. First, this chapter will not take a stance on any of the epistemological debates about self-knowledge. I will assume that we can, in fact, have knowledge of ourselves, and I will remain silent on the question of whether this knowledge is of a fundamentally different kind than other types of knowledge. Thus, I say nothing about the nature or extent of our so-called ‘privileged access.’ The points I want to make here concern what we are morally required to value for its own sake and do not concern the means by which we can achieve what we are morally required to value.

Second, as I noted in the introduction, my aim is not to show that we ought to value all self-knowledge non-instrumentally. There are many trivial facts about ourselves—like the number of hairs on our head, the exact distance between us and the North Pole, or which sock we put on first this morning—that are at most valuable instrumentally. Rather, this chapter will show that a specific kind of self-knowledge, which Quassim Cassam has called ‘substantial self-knowledge,’ should be non-instrumentally valued. Substantial self-knowledge, according to Cassam, is self-knowledge that pertains to facts about what makes us happy, what we value, what our characters are like, where our abilities and aptitudes lie,

---

96 Cassam, Self-Knowledge, 29. While I agree with Cassam that substantial self-knowledge is important to humans in a way that more trivial self-knowledge is not, I disagree with his claim that substantial self-knowledge is still ultimately instrumentally valuable. Rather, as I will argue in Sections II and III, self-knowledge is something that we have reason to value for its own sake.
how we feel emotionally, and what we believe.\textsuperscript{97} While Cassam provides criteria for determining whether a particular fact about oneself qualifies as self-knowledge, he does not explain why these criteria are the right criteria or give an account of what, if anything, ultimately unites substantial self-knowledge into a unified and distinctive category.\textsuperscript{98} As I have suggested in the Introduction, there is something that fundamentally unites all of the types of self-knowledge that Cassam labels ‘substantial’. Specifically, the types of self-knowledge that count as substantial are those types that pertain most directly to our identity-constituting ends. At the end of this chapter, I will explain how my account sheds light on why the category of substantial self-knowledge takes this particular form.

\textbf{Section 2: Love’s Reasons}

Before we ask what reasons we have to know ourselves, let’s first ask another question: what reasons do we have to know other people? Obviously, we have some purely pragmatic reasons. If you want a promotion at work, you will have a reason to learn whether your boss is a stickler about deadlines. If you want good relations with your neighbors, you will have a reason to find out how tolerant they are of loud parties and unkempt lawns. 

We also have moral reasons to know people. To treat you with respect requires me to know that you’re a person, insofar as it is your status as a person that largely determines what

\textsuperscript{97} Ibid, 30. Cassam does not include life experiences in his description of substantial self-knowledge, but for reasons that I will make clear in Section IV, I think that this omission is a mistake.

\textsuperscript{98} While Cassam, \textit{Self-Knowledge}, 36–40, does offer ten conditions that most, if not all, substantial self-knowledge purportedly meets, he refrains from offering any sort of unifying justification of these conditions, and he is careful to note that the items on his list are not necessary and sufficient conditions for substantial self-knowledge.
it means for me to act respectfully towards you.99 But knowing that someone is a person
gives us no insight into who she is as an individual person, and it is this sort of knowledge
that interests me presently.

At times, morality does give us reason to obtain this sort of individualized
knowledge. To treat people with respect may require that we be cognizant of what systems of
etiquette they observe, or of what their life is like. The value of this individualized
knowledge, once again, is instrumental—a little bit of personal knowledge about the people
with whom we regularly interact helps ensure that we treat them respectfully in accordance
with their own values. To respect your coworker, for example, requires that you know
whether he observes any religious or ethical dietary laws in order to avoid bringing a dish to
his dinner party that he finds offensive.

We have plenty of instrumental reasons to know other people. But we have reasons in
addition to these to know the people with whom we share close friendships, familial
relationships and romantic partnerships. These types of relationships, which I will call
‘loving relationships,’ are partially constituted by their participants’ love for each other. To
have a friendship with someone, for instance, is in part to love that person as a friend. And
love provides us with reasons that cannot all be instrumental in nature. To understand what
these reasons are, let’s consider the distinctive mode of valuing that love involves. J. David
Velleman describes loving people as caring for them as ends in themselves.100 We may
appreciate the barista as the means to our morning coffee, but when we love someone, we do


so without needing to appeal to the instrumental role that he or she plays in any of our ends.

Harry Frankfurt describes love in similarly Kantian terms:

When we love something…We care about it not merely as a means, but as an end. It is in the nature of loving that we consider its objects to be valuable in themselves and to be important for their own sakes.\(^{101}\)

When we love people for their own sakes, our love for them becomes a source of reasons. As Bernard Williams’ famous example of the drowning wife demonstrates, our reasons for aiding, preferring, and supporting our loved ones need not bottom out in considerations that are in any way independent of the fact that you love them. Instead, love gives us the reasons that inspire our devotion to our loved ones.\(^{102}\) These reasons must be at least sometimes non-instrumental. To illustrate, consider the reasons you have to adopt your loved one’s ends as your own. Perhaps you do so to make her happy or to help her live a fulfilling life. But even if your reasons for adopting her ends are instrumental, the reasons that you have to care about her happiness or self-fulfillment are not: your concern for these things can be fully explained by the love that you have for her.

To be sure, loving people typically involves more than just caring about them for their own sake. In particular, love typically involves specific desires, emotional experiences, vulnerabilities and so forth. But the particular desires and emotions that we feel towards our

---

\(^{101}\) Harry Frankfurt, “On Love, and Its Reasons,” in *The Reasons of Love* (Princeton: Princeton University Press, 2004), 42. Velleman and Frankfurt disagree about what interpersonal love is responsive to. Velleman argues that interpersonal love, like interpersonal respect, is responsive to an agent’s intrinsically valuable *personhood*—the love that we have for a person’s quirks and character traits is thus a “response” to those traits as “an expression or symbol or reminder of his value as a person” (106). Frankfurt denies this, and instead argues that love need not involve an awareness of an object’s intrinsic value. It is worth noting that Frankfurt works with a far broader conception of love than Velleman: while Velleman focuses his attention on romantic love, Frankfurt is interested in the way in which we love romantic partners, friends, traditions, scientific truth and so forth. “On Love,” 42.

\(^{102}\) It might alternatively be said that love *is* this devotion or that love consists in having certain types of reasons. For clarity’s sake in this chapter, I will talk about love as giving us particular types of reasons.
loved ones will vary greatly depending on the type of love that we have for them. In contrast, as Susan Wolf observes, caring about a person for their own sake is a feature of love that is common to every type of loving relationship.¹⁰³

Much of the philosophical literature on love’s reasons focuses specifically on reasons for action. Williams’ example of the reason that the man has to save his drowning wife over another drowning person—‘Because she’s my wife’—is the paradigmatic example of this tendency. These non-instrumental reasons for action are commonly understood to be an essential part of what it is to love another person: if you don’t see yourself as having reason to help the object of your affection out for her own sake, you don’t love her.

In addition to providing us with reasons for action, love also gives us epistemic reasons—reasons to pay attention, to listen, and to try to know. To understand the pervasiveness of love’s epistemic reasons, think back to the fascination you felt towards your first crush or the interest you took in learning about your grandfather’s war stories. If I were to ask you to explain why you were interested in his war stories over those of the other veterans in the nursing home, it would be sufficient for you to say ‘I’m interested because I love him’. Loving someone makes them interesting to us, and in particular gives us a reason to know them over and above any pragmatic or moral reasons we may have already had.

One might object that, in talking about the non-instrumental reasons that love gives us to know others, I have stacked the deck in my favor by appealing to the beginnings of love, rather than its middle or end. To be sure, the fascination that we feel for our loved ones often dissipates as we accomplish our goal of knowing them. It is easy to listen with rapt attention

to everything that your first crush tells you about himself, but it is much more difficult to pay the same attention to a longtime friend and partner. The diminishment of interest that we sometimes feel towards our longtime loves is, however, simply what happens when we have largely accomplished the end that this particular reason of love sets for us. Seeing someone as worth knowing for the sake of knowing is perfectly consistent with not finding the story of the time she got stuck in an elevator as fascinating on the twentieth telling as it was on the first. Even in these long term loving relationships, however, we still have non-instrumental reasons to know. This becomes clear when we look at one of the hallmarks of falling out of love with someone: when, after a time, we lose interest in learning more about our loved one, our disinterest is usually a sign that our love has run its course.

Of course, we can love people with whom we do not share loving relationships. Outside of the context of loving relationships, love’s reasons can be normatively weak. The senior partner’s love for her junior colleague might, for instance, give her reason to learn about his life story, but she would owe him no explanation if she chose to privilege the reasons she has to maintain professional decorum over the reasons love gives her to get to know him. But when one is in a loving relationship with another person, love’s reasons become normative requirements.\textsuperscript{104} By ‘normative requirement,’ I do not mean to suggest that the reasons that love gives within the context of loving relationships can never be overruled. Even in our most intimate loving relationships, we regularly and reasonably privilege other considerations over the reasons that love gives us. And in most cases, we cease to be bound by these normative requirements by choosing to end the relationship that

\textsuperscript{104} The normative demands of loving relationships are not all demands of love. Marriages, friendships, and familiar relationships are complex social institutions, the normative requirements of which can be shaped by social convention and legal statute.
gave rise to them. What I mean to suggest by the phrase ‘normative requirement’ is that, when we are in a loving relationship with another person, we typically owe that person a justification when we choose to privilege competing interests over her, and that our unwillingness to act on the reasons that our love gives us can at times make us a proper object of censure. If I decide to go out for drinks with colleagues after work instead of commiserating with my newly-heartbroken best friend, then she could rightfully demand that I explain myself to her. In that particular case, my explanation needn’t be particularly strong to be sufficient—it would be enough to point out that I can’t be expected to commiserate with her all the time. In contrast, were I to choose to go drinking with colleagues instead of loitering around the coffee shop where my unrequited crush always studies, I would not owe him any sort of explanation for my absence. My love for my crush might give me reason to frequent his favorite coffee shop, but it in no way obligates me to do so.

Among the requirements that loving relationships put on their participants, I believe, is a requirement to know each other. Why is this? To answer, it is helpful to examine a relationship whose participants are not committed to knowing each other. Consider the scene in Alice Munro’s *The Beggar Maid* in which the protagonist, Rose, considers why her soon-to-be husband Patrick loves her:

> Patrick loved her. What did he love? Not her accent, which he was trying hard to alter, though she was often mutinous and unreasonable, declaring in the face of all evidence that she did not have a country accent, everybody talked the way she did. Not her jittery sexual boldness (his relief at her virginity matched hers at his competence). She could make him flinch at a vulgar word, a drawling tone. All the time, moving and speaking, she was destroying herself for him, yet he looked right through her, through all the distractions she was creating, and loved some obedient image that she herself could not see.105

---

Patrick loves Rose, in a way. But what makes his love for her, and their resulting relationship, unideal is that he is fundamentally uninterested in discovering more about the person whom he loves. Since the real Rose makes Patrick uncomfortable, he opts to hide her behind a more palatable image of low-class gentility. When he meets Rose’s stepmother, the only mother Rose has ever known, Patrick takes great pains to insist that her “real parents can’t have been like that”:

[Rose] saw that he was trying to provide for her a more genteel background, perhaps something like the homes of his poor friends: a few books about, a tea tray, and mended linen, worn good taste: proud, tired, educated people.\(^{106}\)

But Rose does no better by Patrick. So uninterested is she with the details of his life that she has to be told by her landlady that he comes from money. Rose marries Patrick, ultimately, and quickly loses the small amount of curiosity she ever extended to him. When she later confesses to having an affair with another man (whom she does, as a matter of fact, find worth knowing), Rose finds herself completely uninterested in Patrick’s reaction to the news.\(^{107}\)

What can Patrick and Rose’s unhealthy marriage tell us about the importance of interpersonal knowledge to loving relationships? First, their case highlights the fact that wanting to know the people with whom we share loving relationships is part of what it means to be in these relationships. This is why, although they had a legal marriage, Rose would not be speaking entirely figuratively if she were to say that there was a sense in which they were never truly married to each other.

\(^{106}\) Ibid, 91.

\(^{107}\) Ibid, 134.
It is, of course, not enough that we simply want to know the people with whom we share loving relationships. Rather, the norms of these relationships require that we actually endeavor to gain knowledge of our loved ones. In this way, we are normatively required to try to know the people with whom we share loving relationships. More often than not, we are happy to fulfill this particular normative requirement. The love that we have for the people with whom we share these relationships, after all, makes us want to know them for the sake of knowing them. At other times, we may find fulfilling this commitment onerous: it can sometimes be frustrating, boring, or even upsetting to learn more about the people whom we love. And yet, when we continue to listen, we demonstrate that we see them as inherently worth knowing, even at the expense of our time, energy or emotional wellbeing.

That love gives us non-instrumental reasons to know our loved ones, and that these reasons become normative requirements within loving relationships, helps explain why relationships based on projection, like Patrick’s relationship with Rose, are so far from ideal. We sometimes fall in love not with people, but with idealizations of people. When we do this, we construct our own story about who they are, and what they value. But when you start a relationship with someone (even if it is your idealization of her that motivates you to start it), you come to owe it to her to appreciate her as she truly is, and not simply as you want her to be.\textsuperscript{108} If you try to keep the relationship alive in the ways that Patrick did, by refusing to listen to your beloved when she tells you things that contradict your projection of her, or by ‘correcting’ her when she expresses an opinion that you want her not to have, you not only

\textsuperscript{108} The particular type of loving relationship that we share with another person also partially determines the types of interpersonal knowledge that will be most interesting and important to us. I want to know both my best friend and my mother for the sake of knowing them, but I don’t want to know the same things about my mother that I want to know about my friend. My mother is, after all, my mother. Given the nature of our relationship, there are certain facts that I do not want to know about her. I’m sure she feels exactly the same way.
threaten the stability of your relationship, but you also mistreat the person with whom you share it.

As an explanation of why we ought to know our loved ones, an instrumentalist story is clearly inadequate. A certain amount of interest in knowing your loved ones is a constitutive requirement of being in loving relationships with them. And a commitment to actually seeking knowledge of our loved ones is something that we are normatively required to have so long as we remain in loving relationships with them. While we may be able to construct a second-order account about the value of this sort of commitment—perhaps it could be argued that a commitment to knowing your loved ones for the sake of knowing them has benefits down the road that can’t be secured by a commitment to knowing them for instrumental reasons—this account would be motivationally inert and alien to our experiences of love and loving relationships.

Section 3: Self-Love’s Reasons

In the interpersonal case, love gives us non-instrumental reasons to know the people whom we love, and being in a loving relationship can obligate us to act upon those reasons. This means that knowledge of the people with whom we share loving relationships is something that we should want for its own sake. But how does this help show that self-knowledge is also something that we should want for its own sake? What, in other words, do the duties of loving relationships have to do with the duties that we have towards ourselves?

109 In his second argument against Cassam’s claims about the instrumental value of self-knowledge, Eric Schwitzgebel also appeals to analogy between interpersonal and intrapersonal knowledge. His argument takes the following form: Knowledge of other people can be intrinsically valuable. Since it would be strange if interpersonal knowledge had intrinsic value, but intrapersonal knowledge did not, we have no good grounds for thinking that self-knowledge lacks intrinsic value. “Self-Knowledge,” 2. We thus have no good grounds for thinking that self-knowledge lacks intrinsic value. While I am sympathetic to this argument, I’m ultimately making a different claim about the non-instrumental reasons we have to care about knowing ourselves.
My answer, in short, is quite a lot. Just as we can be in better and worse interpersonal relationships, so too can we relate in better and worse ways to ourselves. Loving interpersonal relationships, I argue, provide a natural model through which to understand the requirements that we must meet in order to relate well to ourselves. When we understand what it is to relate properly to ourselves, we will be able to see why we ought to be committed to knowing ourselves for the sake of knowing ourselves.

What It Means to Relate to Ourselves

What does it mean to relate to oneself? Initially, the language of ‘intrapersonal relationships’ might be off-putting to some. Paradigmatically, we think of relationships as involving two people. If this were a necessary feature of relationships, then it would not be possible for us to talk about relationships with ourselves without first assuming some dualistic theory of selfhood. But the very fact that we understand notions like self-love, self-hatred and self-knowledge shows that relationships need not always involve two people. Rather, it is a part of ordinary life that we do have a relationship with ourselves. Here, I am not making any metaphysical point. By a ‘relationship with ourselves, I mean merely to point out something that is assumed in every form of self-regard, be it cognitive or emotional: viz. that we are capable of simultaneously regarding ourselves and being regarded. And the relationship that we have with ourselves, like the relationships that we have with others, can assume better and worse forms. We can loathe or love ourselves, and treat ourselves well or poorly.

As such, it is worth asking what it means to be in a healthy intrapersonal relationship. Joel Feinberg’s discussion of self-love in Absurd Self-Fulfillment gives us the beginnings of an answer:
… self-identity can be conceived as a kind of arranged marriage (I did not select the self that was to be me) that in a stable person ripens into true love, but in an unstable one sours into rancor and self-destruction.  

It is telling that in describing what it means to have a healthy self-identity, Feinberg draws an analogy not with just any interpersonal relationship, but with marriage in particular. Marriages, when they’re going well, are paradigmatic examples of loving relationships. Nor is Feinberg alone in analogizing the way in which we ought to relate to ourselves to a loving interpersonal relationship. This type of analogy goes back in philosophical history at least as far as Aristotle, who described the relationship that the properly self-loving person has with himself as a friendship. And it is also familiar in our non-philosophical discourse. Fashion mogul Diane Von Furstenberg, for instance, once began a speech to a class of graduating high school girls with the recommendation that they strive to become their best friends, while playwright Oscar Wilde likened self-love to “a life-long romance.”

The basic thought contained in all of these descriptors is that our relationship with ourselves, which is necessarily intimate, must be loving in order to be healthy. Why is this? Certainly, most types of relationships can be healthy even if their participants feel neutral towards each other. I don’t have to care about my accountant as a person, for instance, to

---

10 Joel Feinberg. “Absurd Self-Fulfillment,” in The Meaning of Life: A Reader, eds. E.D. Klemke and Steven M. Cahn, (New York: Oxford University Press, 2008), 178. See also 177, where Feinberg notes that loving one’s true nature “for better or worse” calls to mind “the wedding ceremony and its conception of marital love as loyalty and devotion without conditional reservation.”

11 Aristotle, Nicomachean Ethics, 1166a1-b29.


13 Oscar Wilde, “Phrases and Philosophies for the Use of the Young,” The Chameleon (1894).
have a productive business relationship with him. And given the nature of our relationship, it would not be wrong of me if I never make any effort to form an attachment to him or his projects. But I would be doing something wrong if I was similarly neutral towards my mother and her projects. She has been in my life since day one, and has shaped the person I am today in too many ways to count. An inability to feel strongly attached to her or an unwillingness to take an interest in her for her own sake would not be tantamount to respectful disengagement, as it was in the case of my accountant. Rather, it would be a sign that I disvalue her and the relationship that we share.

In much the same way, I am too close to myself to remain neutral on the subject. I must have some sort of attitude towards myself. Intuitively, a healthy relationship with myself won’t involve a negative attitude. Rather, I think that Feinberg and Aristotle are right to suggest that properly relating to oneself requires self-love. And so, it is to self-love that we will now turn.

**Why Our Relationship with Ourselves Demands Self-Love**

Self-love, like its interpersonal analogue, is a mode of non-instrumental valuing. The reasons we have in virtue of loving ourselves, like the reasons we have in virtue of loving other people, are often non-instrumental reasons. We can see these non-instrumental reasons clearly when we consider why we are invested in advancing our own interests in a way that we are not invested in advancing other peoples’ interests. The frazzled assistant professor might be able to answer the question of why she cares more about her tenure file than her colleague’s by pointing to the instrumental connection between getting tenure and

---

114 Self-love, like interpersonal love, involves attitudes, emotions, and desires. For my present purposes, it is sufficient to focus on one aspect of self-love, *viz.* the non-instrumental reasons that self-love gives to know ourselves.
increasing her well-being. But she would be hard-pressed to answer the further question of why she is uniquely invested in her well-being in similarly instrumentalist terms. Rather, her well-being is uniquely important to her for the same type of reason that it is uniquely important to her brother. Just as her brother loves her, so too does she love herself, and that love gives both of them non-instrumental reasons to be invested in her well-being.

In addition to giving us reasons to be especially invested in our personal projects, self-love also supplies us with non-instrumental reasons to know ourselves. That we pay a level of attention to ourselves that would be inappropriate if directed at almost anyone else speaks to this fact. Our attempts at figuring out our motivations, interests and commitments are often both driven by, and expressions of, our self-love.

This explains why an absence of self-love is often accompanied by an indifference towards self-knowledge. Think about some of the people whom we would intuitively consider to lack self-love. On one side of the spectrum, you will find the cult member who has committed herself to disowning every aspect of her individuality. She is not about to try to puzzle out who she is, as she no longer finds that person worthy of attention. On the other side, you’ll find people who loathe themselves, or who have crippling low self-esteem. To see how these forms of self-regard frustrate our ability to value self-knowledge, consider the case of ‘Suzanne’, a woman who struggled to have a sense of self-worth after being forced as a teenager to give a baby up for adoption. When asked to share her story with Ann Fessler,

---


who was writing a book on the subject of adoption in the decades before *Roe v. Wade*, ‘Suzanne’ hesitated. “There’s still that voice in me that says, “Who would be interested? No one cared then, why would they care now?” she remarked, “I was abandoned when it was right in everybody’s face, so I still believe that nobody cares. My personal struggle is to get beyond thinking I’m not worth caring about.” She said. ‘Suzanne’s’ struggle to lovingly care for herself is not separate from her struggle to see her own story as worth retelling. Rather, her inability to love herself is partially constituted by the inability to see her life story as deserving of attention.

To be sure, self-loathers are sometimes prone to paying a great deal of attention to themselves. We can, for instance, imagine someone whose self-hatred motivates him to endlessly fixate on his flaws. In these cases, I think self-discovery can often function as self-flagellation: the self-loather looks at himself not because he sees himself as worth knowing for the sake of knowing, but because he hates the person he sees and thus wants to punish that person by forcing himself to take a closer look at an image that he finds repugnant.

All that we need to establish the claim that self-knowledge is worth having for its own sake is the fact that self-love gives us the same non-instrumental reasons to know ourselves as interpersonal love gives us to know our loved ones. But I want to establish a stronger claim, viz. that we *ought* to seek self-knowledge for its own sake. On its own, self-love cannot give us this claim. If we love ourselves, we will be motivated to know ourselves

---

convention of referring to ‘Suzanne’ in scare-quotes to mark the fact that her name was changed to protect her identity.

117 Ibid, 7.
for the sake of knowing ourselves. But if we do not love ourselves, as ‘Suzanne’s’ case illustrated, we will not be able to see ourselves as worth knowing in this way.

‘Suzanne’s’ case actually gives us the resources that we need to get out of this objection. To see what these resources are, we must keep in mind that, while ‘Suzanne’ may have internalized the message that her teenage pregnancy makes her unworthy of love, she nevertheless continues to want to love herself. She has the same reasons to want self-love that we have to want to love our family members: we’re stuck in close relationships with both our family and ourselves that require love from their participants if they are to be healthy. And so, ‘Suzanne’ will continue to have reason to love herself even if she can’t stand herself.

But here a worry arises. Although most familial relationships require love in order to be healthy, we do not typically think that people can be required to love each other. This is because we have little control, if any, over interpersonal love. At most, we can try to cultivate our love for others by endeavoring to see them in a sympathetic light in the way that Iris Murdoch described.\footnote{Iris Murdoch, “The Idea of Perfection,” in \textit{Existentialists and Mystics: Writings On Philosophy and Literature} (New York, NY; Penguin Books, 1997), 312-313.} But if your unpleasant aunt is, even on the most sympathetic interpretation, truly unlovable in your eyes, then you cannot be blamed for your inability to love her. If love isn’t something that is under our voluntary control, how can we owe it to ourselves to cultivate self-love?

To answer, consider one important difference between the control that ‘Suzanne’ has over herself and the control that we have over our least favorite relative. While you cannot choose to make your insufferable aunt more loveable, ‘Suzanne’ can choose to make herself

into someone whom she can love. And this is what she does when she agrees to do the interview. She goes out on a limb, and affirms something about her value that others have conditioned her to deny. In doing so, she not only treats herself with the partiality that self-love requires, but she also acts in a way that makes her more loveable to herself.

Insofar as we have self-control, we will be able to act in ways that either promote or frustrate self-love. And given the intimate nature of the relationship that we have with ourselves, I think we should strive to love ourselves and make ourselves loveable. Part of what it takes to accomplish this task, as ‘Suzanne’s’ story demonstrated, is a commitment to acting on the reasons that love gives us to know ourselves. This raises a question: what does it mean to act on these reasons? We should not expect a theory of self-love to supply us with an answer. This is because, while self-love can draw our attention inwards, it also has the power to make us blind to parts of ourselves. Shakespeare captures and admonishes the selective attention that self-love often promotes in one of his sonnets:

\[
\text{Sin of self-love possesseth all mine eye} \\
\text{And all my soul, and all my every part;} \\
\text{And for this sin there is no remedy.} \\
\text{It is so grounded inward in my heart.} \\
\text{Methinks no face so gracious is as mine,} \\
\text{No shape so true, no truth of such account;} \\
\text{And for myself mine own worth do define,} \\
\text{As I all other in all worths surmount.}^{119}
\]

If there is something morally admirable about acting on love’s non-instrumental reasons to know oneself, we will not find it in Shakespeare’s narcissistic self-lover. But what exactly is missing from this type of self-inquiry that keeps it from being admirable? One answer that I will ultimately reject is that Shakespeare’s self-lover is, paradoxically, lacking self-love. This

---

is the type of answer that comes out of some prominent modern and ancient philosophical accounts of self-love. Consider, for instance, why Harry Frankfurt criticizes Kant for drawing too close a parallel between self-love and self-indulgence:

As I understand self-love, it is quite unlike the attitude that Kant has in mind when he laments that we hold the self too dear. In speaking of those who love themselves, Kant describes people who are motivated predominantly by an interest in satisfying their own inclinations and desires, and who on any particular occasion will naturally be moved to act by whichever of those inclinations and desires happens to be strongest. Those people are not being driven by what I think of as self-love. Their attachment to the dear self is less like self-love than it is like self-indulgence, and self-indulgence is something else entirely.120

To illustrate this opposition, Frankfurt offers the example of the love that parents have for their children. “Parents who love their children,” he remarks, “take great care, if they are sensible, to avoid being indulgent.”121 While I agree with Frankfurt that sensible loving parents take care to avoid indulging their children, I think he’s wrong to suggest that parents must be sensible to be loving. Indeed, parental love can frustrate sensibility: if they didn’t love their children, I imagine that many parents would find it easier to avoid being indulgent.

For the same reason, we should be suspicious of accounts of self-love that locate its true form in the ideal. Just as loving another person can cause us to overlook their faults, so too can self-love move us to attend to only the prettiest portions of our personhood. Self-love thus only gets us part of the way to explaining how we owe it to ourselves to seek self-knowledge for self-knowledge’s sake. While self-love gives us non-instrumental reasons to know ourselves, it does not tell us how we ought to act on those reasons.


121 Ibid, 79.
Here, it is helpful to contrast the attention that Shakespeare’s self-lover pays to himself with a more ideal form of self-inquiry. As an example of the ideal, let us turn back to Rose, the protagonist in Munro’s *The Beggar Maid*. For all the lack of attention she paid to her husband Patrick, it is clear in the novel that she has the sort of self-insight that many of us strive towards. In one scene, Rose finds herself wondering why she dislikes a younger party hostess. While looking closely at the woman, Rose reveals her own self-awareness:

On the other side of the spool bed was a large mirror, hung suspiciously high, and tilted. Rose tried to get a look at herself when the girl was bent over the basket. It is very hard to look in the mirror when there is another, and particularly a younger, woman in the room. Rose was wearing a flowered cotton dress, a long dress with a tucked bodice and puffed sleeves, which was too short in the waist and too tight in the bust to be comfortable. There was something wrongly youthful or theatrical about it; perhaps she was not slim enough to wear that style. Her reddish-brown hair was dyed at home. Lines ran both ways under her eyes, trapping little diamonds of darkened skin.

Rose knew by now that when she found people affected, as she did this girl… it was usually because she, Rose, hadn’t received and was afraid she wouldn’t receive the attention she wanted, hadn’t penetrated the party, felt that she might be doomed to hang around on the fringes of things, making judgments.

While Rose’s knowledge does not inspire any behavioral transformations—after noting her insecurity, she proceeds to seek out the very same type of attention that she wishes she didn’t crave from the other party-goers—there is nevertheless something to admire about it. Rose looks at herself in the mirror (quite literally) in order to see herself as she is, and not as she wants herself to be. In this way, she extends to herself the same careful attention that we hope our loved ones extend to us.

---

122 In this scene, Rose examines the hostess with the same type of ‘loving attention’ that M. directed towards D. in Iris Murdoch’s classic example (*Beggar Maid*, 312-313).

123 Ibid, 157-158.
Rose’s attempts at self-scrutiny are admirable in a way that Shakespeare’s self-lover’s narcissistic navel-gazing is not. And the difference between Rose and the self-lover, I think, has to do with the seriousness with which each attends to their project of self-discovery. While Rose and the self-lover both love themselves and see themselves as worth knowing for the sake of knowing, only Rose takes seriously the task of knowing herself. And taking the ends we set for ourselves seriously, as demonstrated in the last chapter, is part of what it means to have respect for oneself as a person.

Section 4: From Self-Love to Self-Respect

In the last chapter, I discussed a notion of respect—which Darwall calls ‘recognition respect’—that people are commonly believed to deserve simply because they are people. Because respect is something that we owe to people regardless of what relationships we share with them, it is something that we automatically owe to the people with whom we share loving relationships. To treat our loved ones as they deserve thus requires us to meet special obligations from them that stem from love’s reasons, and also to treat them with the respect that all persons are owed. This is why it is possible for loving relationships to be loving without being healthy: two spouses can love each other deeply, but their marriage will still be destructive if they do not treat each other with respect. Similarly, I cannot be said to be relating properly to myself if I don’t respect myself as a person. Thus, whatever else relating properly to oneself involves, it must necessarily involve self-respect.

124 I will be using ‘recognition respect’ and ‘respect’ synonymously in this section.

125 For a similar take on the relation between the requirements of recognition respect and the requirements of loving relationships, see Scanlon, What We Owe to Each Other (Cambridge: Belknap Press, 1998). While Scanlon uses different language to describe this relation, he does aptly observe that responding appropriately to one’s friends involves not just partiality and special affection, but also the recognition that they have “moral standing as persons, independent of our friendship, which also places limits on our behavior” (165).
In the interpersonal case, treating people with respect primarily involves non-interference. We must, first, refrain from acting towards others in ways that are dehumanizing or degrading, and second, avoid unduly interfering in the decisions that they make for themselves about how to live. To respect you, I shouldn’t humiliate you for my own entertainment, paternalistically meddle in your affairs, or force you to make the life choices that I think are best.

Interpersonal recognition respect, if anything, holds us back from knowing other people. It would be easier for me to get to know you if I could read your emails and diary entries. But the respect that I have for your privacy keeps me from hacking into your Gmail account or stealing your journal. While love can give us reasons to know other people, respect often prohibits us from taking the most direct route to that knowledge.

But while interpersonal respect can hold us back from knowing other people, self-respect does not prevent us from knowing ourselves. It is disrespectful for me to read your diary without your consent because respecting you requires that I abide by your decision to keep certain things private from me. But there is no similar barrier when it comes to my own diary: I wrote it, and so I decide who gets to read it.

Indeed, self-respect actually gives us a positive reason to strive to make sure that our attention to ourselves results in self-knowledge rather than empty self-flattery. This is because, while interpersonal respect often requires that we refrain from interfering in the decisions that other people make about how to live, self-respect carries with it the opposite

---

126 To be sure, love can also at times hold us back from invading our loved ones’ privacy. When we love someone, we are often disposed to please them and to abide by their wishes. But this is not a necessary feature of love; just as love can sometimes motivate us to respect our beloved’s wishes, so too can it sometimes motivate us to overstep our bounds and violate her privacy.
demand. It is a matter of self-respect, Robin Dillon notes, that we aspire to live in a way that
that instantiated our “ideals, aspirations, commitments, and ‘points of no return’.”\textsuperscript{127} The
rationale for this requirement of self-respect is both simple and intuitive: if I am to respect
myself as a person, I must respect the decisions that I make about what kind of person I want
to be, and doing that requires me to follow through with those decisions. If I fail too often to
live in a way that instantiated the decisions that I’ve made about who I want to be, then I fail
to respect those decisions, and by extension the person who made them.

Respecting the decisions we make in life, of course, involves more than simply
following through with them. Specifically, it also involves a willingness to follow through
with them in the right sorts of ways, given the types of decisions they are. To see what I
mean, consider why an athlete might describe her refusal to slack off during practices or her
unwillingness to compete against opponents whom she knows she can easily beat as a matter
of ‘self-respect’. In tying her self-respect to her athletic performance, the athlete is not
pointing out a universal principle to which every rational agent is bound. Rather, she is
drawing attention to the fact that there are better and worse ways to be an athlete. Given the
importance that she places in being an athlete, she recognizes that respecting herself requires
that she strive to live up to certain athletic ideals. When she fails to do this, she fails to take
one of her central commitments, and by extension herself, as seriously as she ought to.

In this regard, self-respect actually demands that we work to overcome some of the
greatest obstacles to self-love. Consider, for instance, the spiral of self-hatred that we can fall

\textsuperscript{127} Dillon, “Lose Your Self-Respect,” 134. See also Dillon, “Self-Forgiveness and Self-Respect,” \textit{Ethics} 112,
University Press, 2009), 133.
into when our acts depart too radically from our values and commitments. The leftist who takes a job at a corporation with a history of human rights abuses in the developing world, for instance, may come to loathe herself for violating her political principles. And with this self-loathing often comes a loss of self-investment: if she’s such a hypocrite, the leftist might think to herself, why should she care about honoring her principles or investing in her personal projects in the future? But self-respect demands exactly this kind of care. To treat ourselves with respect, we must continue to strive for the same ideals that we may have fallen short of in the past. In this way, self-respect demands both that we act in ways that will make us more loveable to ourselves, and also that we be invested in ourselves and our lives in the way that self-loving agents naturally are.

It is here that we can locate the distinction between Rose and Shakespeare’s narcissistic self-lover. Both have been moved by self-love to commit themselves to a project of self-discovery. As with any project, there are standards that one must meet in order to properly pursue self-knowledge. The respectful self-discoverer shouldn’t go too easy on herself in her pursuit of self-awareness. She should be willing to confront harsh truths, and to accept that some parts of her self-conception might need to change in light of new evidence. And this is what Rose does. When she looks at herself with a critical eye, she takes seriously one of her ends. Self-love may have driven her to set self-knowledge as an end, but self-respect propels her to pursue it in the right way. In pursuing self-knowledge in this way, Rose affirms that her ends (of which self-knowledge is one) are important enough to her that

---

128 It is an open question how we should characterize the end that Shakespeare’s self-lover sets for himself. He may have set self-knowledge as an end, which he then failed to realize. Or he may merely believe that he set self-knowledge as an end, when in reality have set something else—selective self-attention, the maintenance of a favorable self-image—as his end.
they’re worth doing right. Shakespeare’s self-lover, by contrast, pursues his project of self-
discovery in a way that shows no similar self-respect. That Shakespeare’s self-lover pays
such selective and biased attention himself suggests that he doubts, on some level at least,
that his actual self is worth the effort it would take to get to know.

In the last chapter of *The Beggar Maid*, Rose recalls a moment in her childhood when
her teacher, fed up with her unwillingness to follow instructions, asked of her “Who do you
think you are?” That this question can be an admonishment to not get too full of oneself,
along with an encouragement towards self-examination,\(^\text{129}\) reflects the extent to which self-
respect, self-love, and self-knowledge are intertwined in our relationship with ourselves.
Self-love moves us to look closely at ourselves, while self-respect arrests our tendency to
fixate on the prettiest parts of our image while ignoring our warts. To take seriously the
project of figuring out who we are, and to see that project as worth doing for its own sake, is
thus not a symptom of narcissistic self-obsession, but rather part of what it is to stand in a
healthy loving relationship with oneself.

**Conclusion: The ‘Substance’ of Substantial Self-Knowledge**

We are now in a position to return to my earlier claim about the unity of the category
of substantial self-knowledge. Recall that in the introduction to this dissertation I noted that
not all self-knowledge is worth having for its own sake. Indeed, the person who thinks that
there is something non-instrumentally valuable about knowing how many of his friends are

---

\(^{129}\) For a discussion of this passage in *The Beggar Maid*, see Per Seyersted, “‘Who Do You Think You Are?’ Alice Munro and the Place of Origin,” *American Studies in Scandinavia* 24 (1992), 21–2. As Seyersted points out, ‘Who Do You Think You Are?’ was actually the original title of the novel. Meanwhile, according to Munro, Rose’s greatest accomplishment in the novel is gaining the ability to answer this formerly titular question: while she does not get “the obvious things, the things she thinks she wants,” Munro notes, Rose does get “a knowledge of herself.” See “What Is” in *For Openers: Conversations with 24 Canadian Writers*, ed. Alan Twigg (Madeira Park, BC: Harbour Publishing, 1981), 19.
more beautiful than him, or how far away he is from the North Pole at any given moment is either condemnable self-absorbed or psychologically bizarre. In contrast, the type of self-knowledge that Cassam calls ‘substantial self-knowledge’ is, I think, something that we have non-instrumental reason to want. We need not appeal to instrumental considerations, for instance, to explain why we are interested in figuring out why we made the major life decisions that we made, or whether we actually value what we purport to value.

What unites these types of self-knowledge into a distinctive category, and grants them their special value? The parts of ourselves that are worth knowing for the sake of knowing, I think, are the parts that are most directly relevant to the relationship that we have with ourselves. And this relationship is a relationship that we have with a person. Persons are, in essence, in the business of setting ends for themselves. Our identity as individual persons, in turn, is shaped by the ends that we have set for ourselves. Given that the reasons that we have for wanting certain types of self-knowledge for their own sake are bound up in what it means to be in a healthy relationship with a person, it is no coincidence that the types of self-knowledge that Cassam labels ‘substantial’ are those types that speak most directly to our personhood.

This gives us the resources to explain why certain types of self-knowledge are not substantial. Just as you can learn your best friend’s blood type without feeling that you have learned anything about her, so too can you learn the exact size of your kidney or that you put our left sock on first this morning without feeling that you know yourself any better as a result. When we discover something about our values, talents, preferences, significant life events or motivations, in contrast, we do in fact learn something meaningful about the ends
that we have set for ourselves over the course of our lives, and thus about the person with whom we share our intrapersonal relationship.

Of course, not every fact about our motivations and life story will be equally interesting to us. What ends we set at a particular time in our life will influence what topics of self-inquiry matter to us most. As an image-obsessed teenager, you were probably more invested in figuring out where your pop-culture loyalties truly lay than you are now. Having a child, meanwhile, can make facts about your own childhood fascinating to you in a new way. A modest person, meanwhile, simply won’t care to know exactly how beautiful she is, or how much her colleagues admire her. But whatever parts of yourself are particularly interesting to you given your life stage and circumstances, you will have a reason to know more about those parts that does not depend on the instrumental benefits that knowing them might bring to you. To be motivated to know yourself for the sake of knowing yourself is part of what it means to love yourself, which in turn is part of what it takes to properly relate to yourself. And properly relating to yourself, by loving yourself and respectfully acting on love’s reasons, is something that we should all strive to do.

130 By delineating the category of substantial self-knowledge, I am able to explain why a non-instrumental commitment to knowing oneself is consistent with virtues of character like modesty and humility. If we modesty as involving either a lack of beliefs about one’s worth or an indifference towards certain types of information about herself, then we can see why a person can simultaneously be modest and non-instrumentally committed to knowing herself. A non-instrumental commitment to knowing oneself for the sake of knowing oneself, after all, does not require that we be interested in knowing everything about ourselves. That being said, my account is inconsistent with Julia Driver’s account of modesty as involving underestimations about one’s self-worth. For Driver, the modest agent is not simply someone who lacks beliefs, or is uninterested in forming beliefs about her greatness. Rather, the modest person is someone who actually falsely believes herself to be less wonderful than she actually is. Driver argues for this account in “The Virtues of Ignorance,” Journal of Philosophy 86 (1989): 373–84 and “Modesty and Ignorance,” Ethics 109 (1999): 827–34. Driver’s interpretation of modesty as a virtue of ignorance has been scrutinized by philosophers such as G.F. Schueler, “Why Modesty is a Virtue,” Ethics 107 (1997): 467–85, Daniel Statman, “Modesty, Pride and Realistic Self-Assessment,” Philosophical Quarterly 42 (1992): 420–38, and Owen Flanagan, “Virtue and Ignorance,” Journal of Philosophy 87 (1990): 420–28.
CHAPTER 3: SELF-DECEPTION AS A PERVERSION OF VALUING

Introduction

In Chapter 2, I developed an account of self-respect that takes into consideration the reasons that we have to respect ourselves as epistemically rational, and not simply practically rational, agents. Self-respect does not simply demand that we refrain from degrading, or letting others degrade, ourselves. Rather, it also demands that we be properly invested in ourselves as rational end-setters. We must, to start, be willing to hold ourselves accountable to standards of epistemic and practical reason: we shouldn’t, in other words, simply form beliefs, but should endeavor to form good beliefs. And we shouldn’t merely set goals for ourselves and act in certain ways; rather, we should care about setting good ends for ourselves. Finally, I argued that self-respect requires self-investment: the self-respecting agent is one who is properly invested in her personal projects, commitments, opinions and so forth. She doesn’t abandon her goals too easily, and she doesn’t go easy on herself when pursuing those goals.

In the last chapter, I located a source of non-instrumental reasons that we have to seek self-knowledge. These reasons are reasons of love. To love something is in part to have reason to want to know it for the sake of knowing it. While loving someone does not by itself demand that we actually act on these reasons, being in a loving relationship with someone—be it ourselves or someone else—can require that we do so. I then explained why self-respect actually demands that we strive to make ourselves into people whom we can find loveable,
and that we extend to ourselves the sort of partiality that is in part definitive of self-love. Self-respect, in other words, requires that we cultivate self-love, which in turn gives us reasons to know ourselves for the sake of knowing ourselves. To set self-knowledge as an ongoing end worth pursuing for its own sake, then, is to respectfully honor one of the commitments that self-love moves us to make. To be invested in going about this project in the right sort of way, in turn, is to affirm one’s status as an epistemically rational agent deserving of epistemic respect from oneself as well as from others.

In this chapter, I consider what, if anything, is so wrong about self-deception. To spoil the punch-line of this chapter—I will conclude, ultimately, that self-deception is a type of epistemic self-disrespect that often involves our identity-constituting ends. In this regard, self-respect involves what I will call a ‘perversion of valuing’ insofar as it simultaneously relates to the ends we value most while also impeding our ability to ever achieve, or even properly engage with, those ends. I will argue, however, that not all cases of self-deception have this feature—some, in fact, can be part of what it takes for us to engage appropriately with the ends we have set for ourselves.

My chapter will proceed as follows. First, I’m going to look closer at the phenomenon of self-deception, and identify six desiderata that I think a successful account of the morality of self-deception ought to meet. Next, I will consider alternate answers to the question of what makes self-deception morally wrong. My aim in this section is not to dismiss any of these alternate accounts, but rather to make clear some of their limits and deficiencies. From there, I will construct my own account of self-deception—the ‘Perverse Valuing’ Account—that meets the desiderata described in section one.
Section 1: The Desiderata of a Successful Account

What does it take to successfully explain the moral dimension of self-deception? To answer this, let’s get a little clearer on what exactly makes self-deception morally interesting.

Desideratum 1: Moral Relevance

First, self-deception strikes most of us as morally relevant in a way that most other failures of self-knowledge are not. Most of the time, we think that there is nothing wrong about having a few false beliefs about what you’re like, forgetting the odd childhood memory, or jumping too quickly to conclusions about why you acted the way you acted. We are, in the end, imperfect epistemic agents with a limited amount of time and a number of competing practical and epistemic ends. Our inevitable failure to have a fully complete, fully accurate set of beliefs about ourselves simply speaks to this fact. And yet, many philosophers through history have often insisted that there is something deeply wrong about being self-deceived. Self-deception, despite involving false beliefs about the self, is according to Butler, Smith, Kant and others, a distinctly moral failing.

Desideratum 2: Moral Uniqueness

Self-deception has often been considered to be a unique sort of moral failing. Pascal, for instance, wrote that while ‘It is no doubt an evil to be full of faults…it is a still greater evil to be full of them and unwilling to recognize them, since this entails the further evil of deliberate self-delusion’. We can generate some thought experiments to sharpen his intuition. Consider first Percy, the unapologetically awful person:

Percy will do almost anything to get what he wants. He lied to his girlfriend about how many previous relationships he had had in order to get her into bed on the first date. He donates to charities, but only when his donations will improve his social standing. He keeps a counterfeit disability parking permit in his car so that he can be sure to always get the ‘choice’ parking spots. He has even repeatedly sabotaged his
coworkers in order to get promoted over them. Percy knows that he’s unscrupulous. When asked to justify his deplorable actions, he’ll tell you ‘I know I’m no saint, but if looking out for myself means I have to be a bad guy, I’m going to be a bad guy.’

Now, Larry:

Like Percy, Larry is a man with hazardously few scruples. He is quick to lie, cheat, steal, and sabotage to get what he wants. Despite his actions, Larry continues to believe that he’s a morally upright person. When faced with evidence that contradicts his self-assessment, he is quick to offer alternate interpretations. ‘Sure, I sabotaged Sally’s work,’ he says, ‘but only because I knew that she was too emotionally fragile to handle the additional responsibilities of a promotion. I did it for her benefit, not mine.’

The difference between Larry and Percy, in short, is that Larry is self-deceived about his moral shortcomings, while Percy is not. Now, nobody would hold Larry or Percy up as moral exemplars. They are, as I have already said, both deplorable. And yet, at least some of us will find Larry to be morally off-putting in a way that Percy is not. Why is this? There is a sense in which the self-deceived are doubly morally condemnable. Percy is morally condemnable because he’s acted wrongly; Larry is morally condemnable both because he’s acted wrongly, and because he doesn’t know it.

Self-deception also seems morally distinguishable from other failures of self-knowledge. Some failures, of course, seem wholly innocuous: we can forget insignificant events that happened to us many years ago, or we can form false beliefs about ourselves based on inaccurate evidence that other people have given us. Other failures seem to be the result of morally-assessable epistemic negligence. Sometimes we jump too quickly to conclusions about ourselves, or are too willing to believe what others tell us about ourselves. A person who routinely fails to know herself for these sorts of reasons may, as I have suggested in Chapter 1, be charged with having set inappropriately low epistemic standards for herself. But even if someone’s self-deception occurs in large part because she has set low epistemic standards for herself, I think it would be a mistake to suggest that her failure to
know herself is morally equivalent to a gullible person’s failure to know herself. At the very least, we typically direct a level of criticism at the former person that would be out of place if directed at the latter.

*Desideratum 3: Moral Content*

Third, the presence of self-deception seems itself to add moral content to non-moral situations. Again, consider two contrasting cases:

When Gloria signed up three months ago to run a 10k with her friends, she was out of shape and undertrained. In the intervening months, she did nothing to rectify either of these impediments: she didn’t train, or try to improve her health. And so, it comes as no surprise to anyone (including Gloria), that on the day of the race, she finds herself huffing and puffing after only running 200 meters, and decides to quit before she reaches the one-kilometer mark.

Now consider Penny,

Like Gloria, Penny was out of shape and undertrained when she agreed to sign up for a 10k race with her friends. And like Gloria, she spent the three months leading up to the race doing nothing to increase her chances of being able to complete it. When the race day is finally upon her, Penny finds herself huffing and puffing through the 200 meters of the route, and decides to bow out before she finishes the first kilometer. Penny’s performance should come as a surprise to nobody, but in fact, it does come as a surprise to Penny. Despite having never run 10k, and despite being generally unfit, Penny nevertheless ardently believed that she’d be able to cross the finish line. Indeed, after the race, she blames anything and anyone but herself for her poor performance. Her shoes were too loose, her laces too tight, she was dehydrated and under-slept, she found the other runners distracting, there should have been more aid stations along the course and so forth. Certainly, she thinks to herself, her poor performance had absolutely nothing to do with her.

Morally speaking, it is hard to see what could possibly be wrong with not finishing a 10k.

And yet, I think that Penny’s self-deception gives her failure some moral content that Gloria’s failure lacks. Her self-deception invites moral judgments. Many of us will likely think that she shouldn’t have had her head so deeply in the sand, and that she ought to acknowledge the real reasons that she failed to complete the race, and that her self-deception reflects badly on her as a person.
Desideratum 4: Range of Reactive Attitudes

In addition to having a negative moral reaction towards Penny, we may also feel a range of non-moral reactive attitudes towards her. Her self-deception may, for instance, provoke feelings of pity and empathy in others. The latter two attitudes can at times be motivational: we may find ourselves wanting to help the self-deceiver maintain her self-deceived beliefs, or we may find ourselves wishing that we could free her of her self-deception. A nuanced account of self-deception’s moral dimension will, in part, involve a sensitivity to the range of attitudes that we can reasonably have towards self-deceivers.

Desideratum 5: Exceptions

Self-deception often has the power to transform morally neutral actions into morally bad ones, and make morally bad actions a little worse. But this is not universally true of self-deception. Indeed, some cases of self-deception seem completely morally innocuous. A fledgling stand-up comedian might deceive herself about just how well her first open mic night really went. In reality, she bombed, but in her own mind she did alright. This self-deception might be what she needs in order to work up the confidence to do more open mic nights, thus sharpening her comedic chops. While the most dogmatic amongst us might still see something wrong with her self-deception, I personally do not. We all, I think, occasionally need to inflate our egos in order to work up the confidence to make valuable life changes and start daunting projects. And there are even cases of self-deception that seem morally beyond reproach. Consider the mother who, in order to maintain her psychological wellbeing while caring for her terminally ill child, retains a self-deceived belief that there is still some possibility that a cure will become available. The situation that this mother finds herself within is so devastating, that it would be horrendous to judge her for her failure to
fully face the facts. A successful account of the morality of self-deception must be nuanced enough to explain why not all instances of the phenomenon are morally on a par.

*Desideratum 6: Coherence with Epistemological Theories of Self-Deception*

Finally, I think that a successful account of the morality of self-deception will either have a means of settling the epistemological questions that arise when we try to conceive of *how* self-deception is possible, or will be able to remain neutral between the two leading answers to this question. In this account, I’m aiming for neutrality. Briefly, there are two philosophical schools of thought about how we can come to be self-deceived. The first school of thought, the Intentionalist Account, models self-deception after interpersonal deception. To be self-deceived is to have some sort of bifurcated self, such that one part of yourself deceives the other part.

The Intentionalist Account is broadly believed to give rise to a few paradoxes regarding how one can possibly convince oneself to believe something one knows to be false, and for this reason some philosophers have abandoned it in favor of a Motivationalist Account of self-deception. According to the Motivationalist, self-deception is the result of biased evidence-gathering and interpreting procedures. A self-deceived agent, Sally, wants \( p \) to be true, collects evidence about the truth-value of \( p \), and then unintentionally interprets that evidence a way that supports her desire for \( p \) to be true. Sally thus ends up believing that \( p \) is true on the basis of her interpretation of that evidence, even though the full body of that evidence better supports the truth of \( \sim P \). I will discuss both accounts of self-deception in greater detail later in this chapter.
Section 2: What’s So Wrong With Self-Deception?

What, if anything, is morally wrong about self-deception? One might first try to answer this question by considering the sorts of effects that self-deception can have on both the self-deceived agent and her moral community at large.

The Autonomy Account

To see what’s so wrong with self-deception, we might first consider the deleterious effect that *interpersonal* deception, broadly construed, is typically believed to have on an agent’s autonomy. When I deceive you about what your circumstances are like, I undermine your ability to make choices for yourself. If I falsely tell you that there’s no job opening at your dream department, you won’t know to apply. If I manipulate you into falsely believing that you have no talent as a singer, you may lose out on the chance to have a fruitful singing career. Similarly, self-deception often frustrates our ability to make informed choices. If Penny can’t see that she is out of shape, she also won’t be able to see herself as having reason to get into shape. In this way, Penny’s self-deceived beliefs serve to restrict the courses of actions that she is able to perceive as being available and desirable.

We must question, however, whether this sort of infringement on autonomy is morally objectionable in the very same way that the infringement caused by interpersonal deception is objectionable. The problem with interpersonal deception, after all, is not simply that it restricts its target’s autonomy. Lots of things do that without being morally problematic: My decision to withhold information from you to which you are not strictly entitled may frustrate your ability to make informed choices about how to act, but it doesn’t follow from this that I have wronged you in any way. The problem with interpersonal deception is rather that it infringes upon an agent’s autonomy in a way to which she could
not possibly consent—if she were to consent to the infringement, after all, she would no longer be deceived. But self-deception does not have the same consent issue. If we assume that we are capable of intentionally deceiving ourselves, then the autonomy objection disappears. If self-deception is something that we intentionally cultivate, as the Intentionalist View suggests, then the affront that it may cause to our autonomy will presumably be morally on a par to the affront that willful ignorance can cause to our autonomy. Since there is nothing in principle wrong about willful ignorance—I can, for instance, permissibly choose not to look at the results of my genetic testing, even if these results would give me important insight into how my life will go—then presumably there is also nothing in principle wrong about self-deception. In both cases, after all, it is me—the autonomous agent—who is ultimately choosing to restrict the potential courses of action available to me in future.

Nor is the autonomy objection particularly strong if one accepts the Motivationalist Account of self-deception. We allow our desires to influence the courses of actions that we perceive as being available to us all the time. Your desire to only date people older than you, for instance, might cause you to overlook some younger, but otherwise promising, suitors. Even if you never realize the effect that this desire is having on your perception of the dating pool, it would be hard to say it is restricting your autonomy in a morally problematic way.

Still, there might be a slippery slope to be found here. Self-deception, as Marcia Baron has pointed out, is often habit forming. To maintain my self-deceived belief that I

---

131Marcia Baron, “What Is Wrong with Self-Deception?” in eds. B.P. McLaughlin and A.O. Rorty Perspectives on Self-Deception (Berkeley: University of California Press, 1988), 437–40. Baron ultimately denies that acts of self-deception are wrong as such, claiming instead that they are wrong either because they serve to shield an
can sing like a songbird, for instance, may require me to also deceive myself into thinking that my electric tuner is malfunctioning, that I didn’t get a part in the musical because the director was intimidated by me, and that the naysayers are all tone-deaf. Eventually, these self-deceptions may radically erode my ability to form reasonable beliefs about the world, and thus to act autonomously. If my view of myself and the world I occupy is too thoroughly warped, I simply won’t be able to make informed decisions about how to act within it. At some point in this process, I will be actively harming the very capacities of my personhood that I ought to be respecting.

Baron is right to categorize extreme cases of self-deception as affronts to an agent’s autonomy. But it would be a mistake to generalize here: most of the time, self-deceptive beliefs restrict our capacity for autonomous action only to an extent which is normally seen as morally permissible. Likewise, while self-deception is often habit-forming, this isn’t necessarily so. Some of our self-deceived beliefs can remain relatively self-contained and temporally isolated: we might adopt them while navigating our way through a period of crisis, and abandon them once the storm has cleared.¹³² And yet, even these sorts of cases can strike many of us as morally problematic. Dale’s self-deceived belief about his wife’s fidelity might not unduly restrict his autonomy, and it might not be part of a larger habit of self-

agent from important responsibilities or because they corrode a person’s belief-formation process, and thus ultimately his agency.

¹³² Sartre would deny this, insisting instead that self-deception is endemic to the human condition, and that self-awareness consists not in the eradication of self-deception, but in the choice to acknowledge its intractability. It is clear that Sartre is labelling as ‘self-deception’ a phenomenon that is very different from the one that interests me at present. To quote Amélie Oksenberg Rorty on Sartrean self-deception: Sartre’s account “is well and good, or at any rate, it is whatever it is. What it cannot be, or purport to be, is an account of those beliefs about oneself that are more specifically self-deceptive than others. If everything one says about oneself is constitutive of a choice rather than a descriptive, questions of truth and validity cannot arise at all. Self-deception, in the narrower sense…disappears altogether.” “Belief and Self-Deception,” 398.
dissembling. And yet, I think most of us feel sorry for him both for the fact that his wife isn’t faithful and for the fact that he doesn’t know it. There’s something bad about being a Dale that transcends the effect that Dale’s self-deception can have on his capacity for autonomous action. If we want a general moral objection to self-deception, then, we will need to look elsewhere.

**The Moral Harm Account**

Perhaps we can object to self-deception by demonstrating how it has the capacity to silence the voice of our conscience. This objection, which I will call the ‘Moral Harm’ objection was most famously developed by Bishop Joseph Butler. On Butler’s charge, self-deception allows us to let ourselves off the hook from the pangs of conscience that typically accompany the performance of an immoral action.\(^{133}\)

There are two ways to develop the Moral Harm Objection. The first draws a connection between self-deception and interpersonal harm: presumably, if you don’t feel guilty for past moral actions, you’re more likely to hurt more people in future. There is something quite plausible about this idea: the pangs of a guilty conscience, after all, often work as a deterrent to future immoral action. The fact that you’d feel terrible if you sabotaged your kind and competent colleague’s chances of promotion, for instance, should typically be enough to dissuade you from starting a slanderous rumor about what she does in the break room after hours. But if you are able to deceive yourself into thinking that such an action isn’t *really wrong*, you’ll have a much easier time starting the rumor.

---

\(^{133}\) Butler, *Sermons*, 113-125.
Let’s call this version of the objection Moral Harm-1. While there is, as I have said, something plausible about it, we should be careful not to overstate the deleterious effect that self-deceivers can have on the world. One ironic feature of self-deception is that its presence is often apparent to everyone but the self-deceiver. Given that self-deception is sometimes so painfully obvious to everybody but the self-deceiver, it can at times actually frustrate the self-deceiver’s ability to cause harm to others. If I know that you’re self-deceived about how callous you really are, I’ll adjust my relationship to you in a way that minimizes the amount of moral harm that you can do to me.

This last point seems to narrow the class of self-deceptions that are genuinely morally harmful: while subtle self-deception carries with an increased possibility of immoral action, overblown self-deception that is obvious to others does not. But if this is true, and if only subtle self-deception is condemnable according to Moral Harm-1, then it will turn out to be the case that self-deception actually has only a small role to play in the objection. Subtle self-deceivers present a moral threat to others not because they’re self-deceived, but because their self-deception isn’t obvious to others. A lack of transparency can be morally threatening even when it isn’t accompanied by self-deception. The conniving politician who is adept at appearing sincere typically poses a greater threat than the conniving politician who isn’t.

---

134 As William Ruddick argues in “Social Self-Deception,” in Perspectives on Self-Deception, eds. B. P. McLaughlin and A. O. Rorty, (Berkeley: University of California Press, 1988) self-deceptions is often sustained through social interactions. Other people can, through concern for our well-being or mere politeness, support the fictions that we have come to believe about ourselves. As well, self-deception can be a joint effort:

Self-deception is simpler when others are similarly engaged. If we have any scruples about our projects, joint action will suppress them, or at least discourage their expression. No one acting in concert has an interest in speaking, or producing evidence, against the false belief or questionable desire that each person wants to maintain. (383)
We might be troubled by the ability that self-deception gives us to silence the voice of conscience for another reason. Conscience is not just valuable as a deterrent to future immoral action. Rather, it can also have a retributive value. The problem with the self-deceiver, then, may not simply be that she is prone to future moral violations, but also that she has not been punished for her current violations. She was able to do the crime, in other words, without having to pay the fine of an angry conscience. Let’s call this Moral Harm-2.

I think there’s something intuitive about Moral Harm-2. We all have uncomfortable truths about ourselves that we’d rather not reckon with, and we may resent self-deceivers for having found a means of not reckoning with them. But the criticism, as it stands, is underdeveloped. To illustrate its incompleteness, let us go back to one of the initial examples: Larry, the man who is self-deceived about the quality of his moral character. We can accuse Larry of ‘doing the crime without paying the fine’: he is, after all, able to pull off any number of moral misdeeds without having to reckon with the fact that he’s a horrible person. But self-deception isn’t the only way that one can avoid having to listen to the chastising voice of conscience. Recall that Percy, Larry’s non-self-deceived analogue, is also able to commit crimes without feeling guilty. But unlike Larry, who accomplishes this through self-deception, Percy avoids the fines simply by not caring about whether or not he’s immoral. If self-deception is wrong, then it isn’t distinctly wrong. Rather, it’s wrong in the very same way that moral apathy can be wrong, víz. that it provides a way for us to act immorally without experiencing any self-reproach.

Note that both versions of the Moral Harm Objection assume a particular picture of the self-deceived agent. As Mike Martin notes, it vividly captures what’s wrong with characters like Albert Speer, Hitler’s minister of armaments and war productions. Speer,
Martin notes, was not a moral monster. Rather, he was a “talented architect and bureaucrat, a loving family man, and considerate to his circle of friends” who nevertheless “managed to remain blind to some of the most horrifying proceedings of the third Reich” and to the role that he played in those proceedings.\textsuperscript{135} But not all self-deceivers are like Speer. Often, we deceive ourselves about things that are morally neutral—our musicality or likeability, how successful we’ll probably be on the job market, or whether we really enjoy the activities we purport to take pleasure in. Sometimes, these morally innocuous self-deceptions might be habit-forming in the way that Baron discussed. And sometimes, they’ll eventually lead people to ignore the call of conscience. But often they won’t: for every one Speer, there are probably thousands of Penny’s who are self-deceived about their looks, habits, and relationships, but not about the morality of their actions.

Given this, if we want either version of the objection to succeed, we’ll have to explain why self-deception can be morally condemnable even when it concerns non-moral matters. This, however, is easier said than done. To go back from another example in this chapter, we can certainly say that Penny, by being self-deceived, has managed to avoid the self-chastisement that most of us perform when we realize how acutely our unhealthy habits have affected our bodies. But we might wonder whether this chastisement is moral chastisement, and also whether it’s chastisement that we really shouldn’t expect or demand that others feel. There might also be a feminist objection that can be raised to this variety of self-chastisement: we might worry that we only expect Penny to chastise herself for being out of

\textsuperscript{135} Mike W. Martin, \textit{Self-Deception and Morality}, (Lawrence: University Press of Kansas, 1986), 38.
shape because we’ve internalized some morally objectionable social norms about how
women should feel about their bodies.

For Moral Harm-2 to work as a description of what is morally wrong with Penny,
we’ll have to be able to show both that self-chastisement in situations such as hers is
appropriate, and also that it’s distinctly moral self-chastisement. But even if we do this, we’ll
run into the same problem that we found in the Larry and Percy case. There are, once again,
many ways in which we can ignore the voice of our conscience. We might imagine a person,
Margie, who is like Penny in a number of relevant ways, but who simply doesn’t care about
being fit. She too gets all the pleasure that comes from never having to go to the gym. And
she also gets to avoid grappling with the unpleasant realization that she’s out of shape simply
by denying that this realization is in fact unpleasant. There might still be something wrong
with avoiding the fines of one’s bad actions in the way described, but I think it’s different
than what’s going on in the case of self-deception. At very least, Margie and Percy are
admirable in a way that Larry and Penny are not. This asymmetry in evaluation, I think, is a
fruitful place to start an account of the wrongness of self-deception. To understand why
Penny’s case is worse than Margie’s, we must first figure out why valuing something about
which one is self-deceived is, in some ways, worse than not valuing that thing at all.

Margie and Penny’s cases bring out something that is missing from both versions of
the Moral Harm Account. While both versions may deal well enough with moralized cases of
self-deception, they struggle to capture non-moralized forms of the phenomenon without
appealing to some version of an improbable slippery slope argument. And neither can explain
why self-deception seems like an especially morally unsavory way to restrict one’s autonomy
or evade self-punishment. I want to suggest here that, when we condemn people like Penny
and Larry, we do so without having to think about the consequences that their self-deception will have on the world, their autonomy, or their moral characters. Their self-deception doesn’t need to lead to pernicious consequences, but is rather quite often in and of itself a moral failing. The question then is: what kind of failing is it?

Section 3: Self-Deception, Self-Respect, Self-Love

Perhaps we can capture the inherent moral wrongness of self-deception by understanding it as a failure of epistemic self-respect. Or perhaps we can say that the self-deceiver has failed herself insofar as she has shown herself to be indifferent, and even hostile, to self-knowledge.

Self-Deception as Epistemic Self-Disrespect

Let us start by developing that first line of argument. As I have previously argued, the reasons that we have to respect ourselves as rational agents extend to our epistemic rationality. And so, just as we ought to care that we’re making sufficiently respectful decisions about how to live, so too should we care that we’re forming our beliefs in a way that is appropriately respectful of our epistemic capacities. If self-respect demands that I not stay in an abusive relationship just because I love the person with whom I share it, so too does it require that I not simply believe things because I want them to be true.

How exactly does the self-deceiver fail to treat herself with epistemic respect? The answer to this question depends on how you conceive of self-deception. I briefly alluded to the two leading theories: the Intentionalist Account and the Motivationalist Account, but I’ll pause now to describe both in greater detail.
The Motivationalist and the Intentionalist

To understand what is going when we deceive ourselves, let us first pause to appreciate what is going when we deceive other people. To begin, interpersonal deception involves at least two people: one who does the deceiving, and one who is deceived. The deceiver in this equation has some belief, \( p \), but intends to lead the deceived to believe \( \sim p \). He accomplishes, or tries to accomplish this goal by saying or doing something that he thinks will cause the deceived to believe \( \sim p \).

Interpersonal deception is not a philosophically mysterious phenomenon. We can quibble over the minutiae—by asking, for instance, whether it’s still interpersonal deception if the deceiver intends only to make the deceived doubt \( p \), or whether we can deceive others through actions, rather than words—but the basic phenomenon is both familiar and straightforward. We understand why people sometimes want to convince others to believe propositions that they themselves believe to be false, and we know what sorts of actions we need to perform in order to get people to believe what we want them to believe.

But whereas interpersonal deception is philosophically straightforward, its intrapersonal counterpart is not. Consider what happens when we try to adapt the description of interpersonal deception given above to account for the phenomenon of self-deception. Suddenly, we have a single agent two both believes \( p \), intends to believe \( \sim p \), and who acts and talks in such a way as to convince himself to believe \( \sim p \). This description of self-deception seems paradoxical for two reasons.\(^{136}\) First, it requires a single agent to

simultaneously believe two contradictory propositions. Such a mental state seems to be impossible, and yet it also seems to be required if we are to understand self-deception as analogous to interpersonal deception.\textsuperscript{137} Second, this model seems to require self-deceivers to perform an impossible task.\textsuperscript{138} If I intentionally undertake a project of deceiving myself, after all, my project would presumably be necessarily self-defeating. In the interpersonal case, my attempt to deceive you into believe $\neg p$ will fail if you find out about my plan. And yet, once again, if we model self-deception after interpersonal deception, then we seem to be forced to accept that this apparently-hopeless project is actually capable of succeeding.

The apparently contradictory nature of self-deception has led some philosophers to deny its very possibility.\textsuperscript{139} More commonly, however, the paradoxes of self-deception have spurred philosophers to action, motivating them to generate new, non-paradoxical accounts of the phenomenon. Roughly, we can partition these accounts into two camps. In the first camp, we will find accounts that seek to evade the paradoxes of self-deception by making it more strongly analogous to interpersonal deception (‘The Intentionalist Account’), while in the second, we will find accounts that seek to show that self-deception is disanalogous to interpersonal deception in some way that renders it non-paradoxical (‘The Motivationalist Account’).

\textsuperscript{137}Mele, \textit{Irrationality}, 121. Mele calls this paradox the ‘Paradox of Belief’.

\textsuperscript{138} Ibid, 138. Mele calls this paradox the ‘Strategy Paradox.’

\textsuperscript{139} For examples of this sort of account, see Stanley Paluch “Self-Deception,” \textit{Inquiry} 10 (1967): 268–78, M.R. Haight \textit{A Study of Self-Deception} (Sussex UK: Harvester Press, 1980) and David Kipp, “On Self-Deception,” \textit{Philosophical Quarterly} 30 (1980): 305–317; Note however that Paluch does not deny that there is some phenomenon that we can reasonably call self-deception. Rather, he denies that it is possible to be self-deceived in precisely the same way that it is possible to be interpersonally deceived (276). In the end, Paluch endorses a model of self-deception that is quite close to the one that Mele outlines two decades later (Mele, \textit{Irrationality}).
Let’s look at the first type of account. Recall that there is nothing paradoxical about interpersonal deception: it is easy to see how two people could hold contradictory beliefs, and how one person, A, could trick another, B, into believing something that A believes or knows to be false. The paradoxes of self-deception both arise when a single agent plays the role of both the deceiver and the deceived. One way out to escape the paradoxes, then, is to strengthen the analogy between self- and interpersonal-deception by positing some sort of divided self. Once we do this, we will no longer have a single agent playing two contradictory roles. Rather, we will have two parts of an agent acting independent of each other to serve as both the deceiver and the deceived.

To evade the paradox requires Intentionalist to posit some internal division within a single self: but which division? Rorty posits a “deep division”, wherein subsections of a single agent are capable of autonomously forming beliefs, experiencing desires, and having intentions. Self-deception, she claims, is aimed at a sort of integration of these selves.140 Herbert Fingarette posits an indeterminate number of primitive “selves” that exist within us through habituated behaviors.141 David Pears contends that there are separate spheres of agency within the self-deceived agent, while Donald Davidson contends that there are “boundaries between parts of the mind.”142

140 Rorty, “Belief and Self-Deception,” 405.


Motivationalist accounts of self-deception have attempted to resolve the apparent paradoxes by weakening the analogy that philosophers have historically drawn between self- and interpersonal-deception. Most typically, philosophers have denied that we intentionally deceive ourselves in the same way that we often intentionally deceive others. This strategy has been the one preferred by Mele, who has described the central cases of self-deception as satisfying four conditions. For an agent’s belief in $p$ to be a self-deceived belief, it must be the case:

1. The belief that $p$ which $S$ acquires is false.
2. $S$’s desiring that $p$ leads $S$ to manipulate (i.e., to treat inappropriately) a datum or data relevant, or at least seemingly relevant, to the truth value of $p$.
3. This manipulation is a cause of $S$’s acquiring the belief that $p$.
4. If, in the causal chain between desire and manipulation or in that between manipulation and belief-acquisition, there are any accidental intermediaries (links), or intermediaries intentionally introduced by another agent, these intermediaries do not make $S$ (significantly) less responsible for acquiring the belief that $p$ than he would otherwise have been.

The ‘manipulation’ mentioned in Conditions 2 and 3 can occur in a variety of ways. Our desire that $p$ may, for instance, lead us to misinterpret as not counting against $p$ evidence that,

---

143 Mele, *Irrationality* is aware that some interpersonal deception is unintentional, and he leaves open the possibility that self-deception may in rare circumstances involve intentionality. His claim, rather, is that ‘garden variety’ cases of self-deception do not involve intention deceiving. He denies, further, that we ought to model self-deception off of unintentional interpersonal deception. “Unintentional deception,” he notes “may be quite accidental. But self-deception seems to be motivated by desires or fears of the agent-patient.” (123).

absent our desire, would plainly appear to us as counting against \( p \). It may also lead us to incorrectly interpret evidence that counts against \( p \) as evidence that actually supports \( p \). And it may lead us to selectively focus our attention on evidence that supports \( p \), while overlooking evidence that supports \( \neg p \).\(^\text{145}\) It is through this desire-driven process of selective evidence gathering and interpretation that we may come to believe that \( p \).

Although the Intentionalists and Motivationalists resolve the paradoxes of self-deception in different ways, we must be careful not to overestimate the gap that exists between them. Both the Intentionalist and the Motivationalist agree that our desires can influence beliefs. But whereas the Motivationalist understands these desires to affect our belief-formation processes in roundabout ways, the Intentionalist contends that they can affect these processes directly. Our desire for \( p \) to be true can lead us to intentionally perform actions that will enable us to form a belief about \( \neg p \)’s falsity.\(^\text{146}\)

Further, although the labels for both accounts suggest otherwise, both Intentionalists and Motivationalists agree that there is room for both intentional and unintentional action within self-deception. For the Intentionalist, the role that intentions play is obvious: when we deceive ourselves, we intentionally cultivate within ourselves a belief that we recognize, in some way, to be unsupported or under-supported by evidence. But once this initial action is created, our self-deceived belief may be maintained through carelessness, apathy, or sheer forgetfulness. For the Motivationalist, the role that intentions play is less direct. Mele, for instance, grants that most cases of self-deception, our desire for \( p \) will motivate us to form

\(^{145}\) Ibid, 125-126.

\(^{146}\) Davidson, “Deception and Division,” 208.
intentions and perform actions related to preserving or acquiring a belief in $p$. What Mele
denies, rather, is that an intention to deceive oneself play a role in all but a fringe subset of
cases of self-deception.$^{147}$

Both accounts are subject to well-known criticisms. The Intentionalist can evade the
paradoxes of self-deception only by endorsing a sharply-divided self that we may find
phenomenologically counter-intuitive. But at the same time that we may raise questions
about the plausibility of this picture, we might also see it as necessary to accurately capturing
the experience of self-deception. The Motivationalist, we might worry, has made the
phenomenon too clean, leaving little to no room for the psychological angst and
indecisiveness that sometimes characterizes self-deception. These concerns, while important
to the continued philosophical exploration of self-deception, are orthogonal to my project in
this chapter. This is because, as I will later demonstrate, my account of the morality of self-
deception can be developed to fit either model of self-deception.

*The Epistemic Failings of Self-Deception*

Let us now return to the question of how self-deception may constitute a failure of
self-respect. I think that both the Intentionalist and Motivationalist can provide an answer to
this question. For the Intentionalist, the failure is direct: the self-deceiver chooses to believe
what she wants to believe, rather than what she has most evidence to believe. For the
Motivationalist, the failure is more subtle: the self-deceiver allows her desires to unduly
impact her evidence-gathering process. In both cases, to maintain her self-deception, the self-
deceiver must lower her standards of evidence-gathering and interpretation with regard to the

$^{147}$ Mele, *Irrationality*, 139.
subject matter of her self-deception. If she is to remain deceived, she can’t go looking too hard for new evidence, can’t ‘play Devil’s Advocate’, and can’t question too deeply the circumstances under which she originally formed the belief.

Most of the time, epistemic negligence is pretty morally unremarkable. When I wake up with a scratchy throat, I may jump to the conclusion that I have a cold. This might turn out to be false—perhaps I have allergies or the flu. And my belief was formed hastily and without much evidence, but the degree of negligence was low, the matter at issue was unrelated to anything that really matters to me, and the false belief that I formed about myself was pretty insignificant. Even willful, desire-directed epistemic negligence can at times seem morally innocuous. The undergraduate’s desire to finish an essay in a course that he loathes may motivate him to perform only the bare minimum amount of research necessary for completing the assignment, thus leading him to take up a position in his paper that he would have found misguided had he done more research. But while his essay writing strategy likely won’t lead him to the honor roll, I also don’t think that it suggests that he disrespects himself as an epistemically rational agent. Rather, it reflects that he, like everyone else, prioritizes certain epistemic projects over others. Just as we can’t be expected to be paradigms of practical reason in every decision that we make about what to do, so too can we not be expected to be epistemic saints when it comes to every belief that we form about ourselves and the world.

If there is something epistemically disrespectful about the self-deceiver’s belief formation, then, we cannot locate it in the simple fact that she is negligent, or even in the fact that she is willfully negligent about forming beliefs. Rather, we must locate her epistemic disrespect in the reasons behind this negligence. One way in which we may uncover these
reasons is by looking at the sorts of desires that seem to motivate us towards self-deception. Mark Johnston has suggested that not all desires give rise to self-deception, but rather that we are motivated towards self-deception by our ‘anxious desires’. Self-deception, for Johnson, is the result of a mental process that aims at the reduction of anxiety, instead of the cultivation of justified true beliefs. Annette Barnes elaborates on this connection between anxiety and self-deception, defining ‘anxious desires’ as follows:

A person has an anxious desire that \( q \) i.e., that it be the case that \( q \), just in case the person both desires that \( q \) and is anxious that it is not the case that \( q \).^{149}

There is, in another words, a bi-directionality in anxious desires: when we experience them, we simultaneously feel a desire that \( p \) obtains, and experience anxiety that \( \sim p \) may end up obtaining. Not all desires have this bi-directionality. We can, Barnes observes, desire a cup of coffee without feeling anxious that we might not get to have it. Self-deception, for both Barnes and Johnston, acts as an analgesic against anxiety. When we deceive ourselves into believing \( p \) we are able, to a large degree at least, put to bed our anxieties over the possibility of \( \sim p \).

We might say then that: the self-deceiver is not epistemically negligent towards just anything, but rather towards a specific class of topics that provoke her anxiety. When it comes to these topics, she is not simply invested in certain propositions being true of them—rather, she is anxious about the possibility that those propositions might not be true. In this way, Barnes and Johnson contend, the self-deceiver can be charged with the vice of

---


epistemic cowardice. The epistemic coward gives priority to safety and comfort over truth and in doing so exercises a morally objectionable level of partiality towards his feelings and desires.

I think that Johnston and Barnes are right to suggest that self-deception is a species of epistemic cowardice insofar as it involves prioritizing something—namely, one’s desires and comfort—over the truth. But this is not to say that all epistemic cowardice involves self-deception. Rather, one can be a clear-eyed epistemic coward. Consider the person who can’t stand to watch the news because she hates to hear about all the atrocities that are going on around the world. Or consider the professor who never opens his teaching evaluations because he can’t stomach the idea of finding out what his students truly think of him. Both of these people can be fully aware that they are cowards when it comes to world news and student evaluations. They can freely admit that, in choosing not to find out specific facts about topics that cause them anxiety, they are privileging their comfort over the truth.

---


151 Barnes, Seeing Through Self-Deception, 166.

152 Here, Barnes distinguishes prima facie objectionability from prima facie moral badness. There is always, she asserts, something prima facie objectionable about epistemic cowardice (and thus, about self-deception), but this does not entail that there is always something intrinsically prima facie morally bad about self-deception. There may be mitigating circumstances that render self-deception objectionable, but not morally bad. The terminally ill cancer victim who deceives himself about his prognosis is epistemically cowardly, and thus objectionable, but it does not follow, Barnes argues, that he is therefore morally bad. While I agree with Barnes that it doesn’t follow from the fact that the terminally ill cancer patient is self-deceived that he is thereby morally bad, I think this is because the very fact that an agent is self-deceived does not allow us to make a judgment about whether they are all things considered, morally bad. We can, however, still make the judgment that their self-deception is morally bad. I admit that I am not clear on where Barnes draws the line between moral objectionability and moral badness. I take it that the distinction is meant to capture the thought that, although there is always something objectionable with self-deception, there may be mitigating circumstances surrounding that self-deception that ought to keep us from actually objecting. In these cases, we can still coherently say that it would be better in some way if the self-deceiver wasn’t self-deceived, while still saying that nobody could reasonably blame him or find him objectionable for being self-deceived (Seeing Through Self-Deception, 163-166).
they can acknowledge that because of this, they’re not ideal epistemic agents. They go too easy on themselves; if they were more courageous, they would be more willing to face potentially hurtful truths about themselves and the world.

Being aware about one’s cowardice does not insulate one from moral censure. We can look down on people who self-consciously shy away from uncomfortable truths in the same way that we look down on people who self-consciously shy away from difficult activities. And our reasons for this censure are manifold. Often, we find ourselves having to pick up the slack for cowards, by taking on tasks that we also dislike, but are nevertheless willing to do. If epistemic cowardice becomes too prevalent, it opens the door to all sorts of human rights abuses. And putting consequences aside, we can condemn cowards for going too easy on themselves, for privileging desires and comforts over rational agency.

Despite this, I think that we have a certain begrudging respect for self-aware epistemic cowards that we do not have for self-deceivers. The self-aware coward is going too easy on herself as an epistemic agent, but unlike the self-deceiver, the self-aware coward is at least not going too easy on herself about the very fact that she’s going too easy on herself. Because of this, there seems to be a moral asymmetry between the self-aware and self-deceived epistemic coward. The questions we must ask, then, are as follows: what does the self-aware coward have that the self-deceiver lacks? Is there something morally objectionable about lacking this thing, whatever it is? And finally, if there is something objectionable, in what way is it objectionable? I think we can make headway towards answering all three questions by exploring the connection between self-deception and self-knowledge.

130
Section 4: Self-Deception and Self-Knowledge

Not all failures of self-knowledge, as I have noted, involve self-deception. What about the reverse? Does self-deception always involve a failure of self-knowledge? In a trivial sense, the answer is clearly yes: the self-deceiver has by necessity failed to know at least one fact about himself, viz. that he is self-deceived. But the topics about which we are self-deceived vary wildly, and clearly go far beyond the domain of the self. To demonstrate, let us consider Stevens, the deeply self-deceived narrator in Kazuo Ishiguro’s *The Remains of the Day*. Stevens takes great pride in having been an exemplary butler to the powerful, proper, and morally upright Lord Darlington.\(^{153}\) And he takes pride in being the son of an equally exceptional butler. In fact, he considers his deceased father to have been a perfect butler.\(^{154}\) But his own recollection of his father’s last days tell a very different story of a man with severely deteriorated physical and mental stamina who had long ago ceased to be great.\(^{155}\) Likewise, although Stevens takes great pains to deny it, he recognizes at some level that his employer, Lord Darlington, was an anti-Semitic Nazi-sympathizer who deserved the public censure that he experienced in his later years.\(^{156}\)

The facts about which he is self-deceived clearly have a close relationship to Stevens. He’s not, after all, self-deceived about the virtues of some other butler’s employer. But while

---


\(^{154}\) Ishiguro, *Remains*, 35.

\(^{155}\) Ibid, 58, 64.

Stevens has a close relation to these facts, it is not a relationship of identity: a fact about Lord Darlington is a fact that concerns Stevens, but it is intuitively not a fact *about* Stevens.157

In one sense, this intuition is right. Stevens is not Lord Darlington, and so facts about Lord Darlington are not in and of themselves facts about Stevens. But if we take a closer look at why Stevens is self-deceived about Lord Darlington, we’ll see that facts about Lord Darlington are, in one important respect, facts about Stevens.

Let’s start by surveying the sorts of explanations we could give as to why Stevens is self-deceived about Lord Darlington’s anti-Semitism. Perhaps Stevens just really hates acknowledging that there are anti-Semites in Britain: he’s proud that his country welcomed Jews fleeing Continental Europe during the Second World War, and the existence of a British anti-Semite frustrates his beliefs about his country’s moral superiority. Or perhaps Stevens just really hates anti-Semites—the very thought that he has *met* one gives him the chills. If either of these explanations were the right one, then we could say that Steven’s self-deception, in one respect, has very little at all to do with Lord Darlington. If Stevens were to find out that it was some other Lord, and not Lord Darlington, who was an anti-Semite, then Stevens would be self-deceived about *that* Lord’s anti-Semitism. But these explanations are clearly not the right one, given the other facts that we know about Steven’s life history. It matters to explaining Steven’s self-deception that we recognize that he had a certain relationship with Lord Darlington, and that this relationship was not merely one of employer to employee. Rather, it was an identity-defining relationship with Stevens: his sense of self

157 Of course, if we conjoin it with other facts about Stevens (e.g. that he is an employer of Lord Darlington), we can derive facts that are directly about Stevens (e.g. that he is employed by an anti-Semite).
was wrapped up with his status as Lord Darlington’s head butler.\textsuperscript{158} Or, to put it another way: the ends that Stevens set in relation to his employment as Lord Darlington’s Butler were identity-constituting ends. Because of this, facts about Lord Darlington were, from Stevens’ perspective at least, also facts about Stevens. In discovering that Lord Darlington was an anti-Semite, Stevens simultaneously learned that he, Stevens, had been employed by an anti-Semite.

In the Introduction, I introduced a distinction between substantial and trivial self-knowledge. I argued in Chapter 2 that we have non-instrumental moral reasons to seek the former, but not the latter sort of self-knowledge. As I have noted, substantial self-knowledge is most typically knowledge relating to the ends with which we, as individuals, most strongly identify. What particular types of facts constitute substantial self-knowledge will vary from person to person. If I identify as an artist, then facts about my artistic abilities will be substantial to me in a way that they would not be to someone who did not identify similarly. Likewise, while most people would not consider the fact that they were once employed by an anti-Semite to count as substantial self-knowledge, Stevens, who identified strongly with his status as Lord Darlington’s butler, would consider the information substantial. And so, when he ultimately realizes that Lord Darlington wasn’t the virtuous man he thought him to, Stevens is forced to grapple with the realization that he’s defined himself in terms of his service to an ignoble Lord.

By the same token, when a husband deceives himself about his wife’s infidelity, he is not only deceiving himself about what she is like, but also about what his life and his marital

\textsuperscript{158} Ibid, 42-43, 126, 138-139, 200-201.
relationship is like. He is, in other words, deceiving himself about facts about his substantial self. Were he to ever overcome his self-deception, he would be forced to reckon not only with the fact that he didn’t really know his wife in an important regard, but also that he didn’t really know certain key facts about himself, his marriage, and his life.

Rorty offers a more vivid and tragic example of the connection between other-regarding self-deceived beliefs and our substantial selves. When King Lear oscillates between acknowledging his dear daughter’s tragic demise, and denying its occurrence, he is, of course, deceiving himself about his daughter. But he is simultaneously deceiving himself about himself:

..of course his identity is at stake: if Cordelia is dead, and he and she cannot tell each other tales and forgive each other, he is a different Lear, a Lear who must bear an identity different from that which could be his, if he and Cordelia’s death were lovingly, playfully reconciled.159

We cannot even understand why Lear is self-deceived about Cordelia’s death unless we appreciate the connection between his relationship with her and his sense of who he is and how his life is gone. Were this connection not to exist—were Lear to be self-deceived about the death of a complete stranger—then his self-deception would be unintelligible to us. And so, Rorty concludes, we employ our strategies of self-deception in the service of maintaining or changing our desired conception of ourselves.160 Or, as Richard Holton pithily put it, self-

159 Rorty, “Belief and Self-Deception,” Inquiry, 400.

deception is “more concerned with the self’s deception about the self, than with the self’s deception by the self.”  

Insofar as we have moral reasons to know our substantial selves, we fail ourselves by being self-deceived. When we deceive ourselves, we privilege our comfort over a clear view of someone who is, to us at least, worth knowing for the sake of knowing. In this way, we can distinguish the self-aware coward from the self-deceiver. The professor who can’t stand to read his teaching evaluations knows that he’s someone who cares a great deal about what others think of him: if he didn’t, he wouldn’t care if his teaching reviews were bad. And he knows further, that in this respect, he’s cowardly: he’s letting his fears of disapproval keep him from reading some teaching reviews that might actually make him happy, or at least might be helpfully instructive. We can say that he’s going too easy on himself, and that he shouldn’t care so much about what people think about him. But what we can’t say is that he doesn’t know himself. While he is cowardly when it comes to reading reviews, he isn’t cowardly when it comes to seeing himself and his flaws.

The self-deceiver, in contrast, suffers from both first- and second-order cowardice. Like the self-aware coward, he shies away from evidence and situations that may cause him anxiety. But unlike the self-aware coward, he also shies away from his cowardice: he is a coward, in other words, not just about specific desires, but also about the nature of his substantial self. Now we must ask: what’s so bad about being a second-order coward?

---

Section 5: Self-Deception as a Perversion of Valuing

Let us return to the second-order cowardice that the self-deceiver exhibits. This is a cowardice directed at the substantial self: it reveals an agent’s unwillingness to recognize herself for who she really is. I’ve already talked about some of the instrumental harms that can follow from this failure of self-recognition, as well as from the failure of epistemic self-respect that it necessarily involves. I’ve argued that instrumental explanations don’t capture the sense in which self-deception can be condemnable for its own sake, while explanations that appeal to the epistemic vices associated with self-deception are unable to distinguish it from other, less condemnable epistemic failings.

In this section, I’ll consider how the self-deceiver wrongs herself through being self-deceived. In doing so, I’ll make clear why it is that self-deception arouses not simply moral umbrage, but also pity. On my account, self-deception often represents a ‘perversion of valuing’ insofar as it both reveals to others what we most value and how we see our substantial selves in relation to those values, while at the same time frustrating our ability to ever actually engage appropriately with those values. In this way, self-deception is both crime and punishment: when we engage in it, we act epistemically carelessly towards the very parts of our lives that most demand our attention, and in doing so, cut ourselves off (often permanently) from those parts.

Taking Our Ends Seriously

To start, let’s recall that a requirement of self-respect is that we be willing to invest in ourselves as end-setters. This involves a willingness to be reasonably partial to the ends that we set, as well as a commitment to not going too easy on ourselves with regards to these ends. Of course, not all ends are created equal: some of the ends that we set matter very little
to us, while others matter a great deal. The ends that matter a great deal are, by and large, the ones that reflect our substantial selves: we set them because they accord with our values, reflect our talents, and shape our lives in the way we want our lives to be shaped. They are, in other words, identity-constituting ends. Information about these ends, and the way in which we are pursuing them, thus qualifies as substantial self-knowledge: if I really care about being an boxer, then I’ll have reason to know whether I won the match because the game was fixed not simply because such information will help me improve as a boxer, but also just because it will tell me something about my substantial self.

What impact does self-deception have on these ends? To start answering both of these questions, let’s look at a famous case of self-deception—this is the case of Norma Desmond from Billy Wilder’s film *Sunset Boulevard*:

The last film that Norma starred in may have been a silent film, but that hasn’t diminished her status as one of Hollywood’s greatest stars. Her adoring fans, whose letters she still receives daily, long for her return to the silver screen. And, after decades away, Norma is finally ready for her comeback—no, not her comeback—her return. She will be playing the heroine in a Hollywood epic that she wrote herself. Paramount loved her script, of course, and De Mille is all set to direct it.

Or at least, that’s how Norma Desmond sees things. The reality of her situation looks rather different. Joe, the scriptwriter whom Norma has hired to ‘polish up’ her screenplay, can’t make heads or tails of her convoluted prose. Mr. De Mille won’t return her calls. And her fan letters—which are, in fact, dutifully written by her former husband-turned-butler—suspiciously all bear the same postmark. Norma may be a Hollywood idol in her own mind, but in the real world, she’s a washed-up has-been with a fragile ego and a grip on reality that’s growing ever more tenuous. Norma’s self-deception tells us something about what she values—but what? It isn’t the case that Norma simply values believing herself to be a famous Hollywood actress. And she certainly doesn’t value being deceived into thinking that she’s a famous Hollywood actress.

---

Rather, Norma’s self-deception reveals her to be someone who deeply values *being* a famous Hollywood actress. Her status as a member of the Hollywood elite, in other words, is an integral part of Norma’s substantial self.

The problem, of course, is that Norma *isn’t* a member of the Hollywood elite: the person she is and the person she believes herself to be have come radically apart as a result of her self-deception. And because of this, she does not merely fail to value her actually self, but expressly *dis*values it. Norma shows that she would rather engage in self-subterfuge than face herself as she truly is.

Norma’s self-deception also tells us something about how she has set ends for herself in relation to one of the things she most values. Here’s where my account diverges depending on whether one has Motivationalist or Intentionalist leanings. I’ll develop the Intentionalist reading first. Recall that Intentionalists think that self-deception is a willful act: to deceive myself requires that I somehow convince myself that $p$ is true, even though I know that it is not. If you’re an Intentionalist, then you should treat Norma’s self-deception as a choice that she has made about how to deal with her loss of celebrity.

But while self-deception might be one way in which we might accomplish (or at least, think ourselves to have accomplished) our ends, it is not the right way. Why not? Why can’t I accomplish my ends simply by thinking that I’ve accomplished them? Why can’t Norma pursue the end that she has set for herself simply by changing her perception of the world? After all, there are times when we *can* do just this. A depressed person might set ‘being happier’ as an end, and she may make progress towards this end by trying to convince herself that her mood has already improved. But the perceptual distortions of self-deception seem relevantly different than the perceptual distortions that the depressed person undertakes. This
is because we don’t get to choose what sorts of actions and behaviors constitute ‘proper engagement’ with our ends. Rather, the behaviors and actions required for ‘proper engagement’ are largely determined by background social norms that we can’t control. Norma can’t regain her fame by distorting her self-perception for the same reason that I can’t become a great baseball player by changing the norms that govern what it takes to play baseball well. The norms of being famous prevent us from achieving fame through self-deception, and the norms associated with being a great baseball player prevent us from simply changing the rules of baseball to suit our talents or lack thereof.

So what would Norma have to do in order to properly pursue this particular identity-constituting end? As a twenty-five year old, Norma knew the answer to this question. She had to accept mediocre roles, keep up her looks, and make nice with directors, producers and film crews. Fifty year old Norma is a different story. For one thing, she simply can’t do some of the things that it takes to make it in Hollywood. She can’t look twenty-five and her anti-social tendencies make schmoozing difficult for her. For another thing, she’s simply unwilling to do some of the things that she’d need to do to make a comeback. She has an artistic objection to talkies and, while she’s too old to play a starlet, she’s unwilling to play a mother.

It seems likely, then, that there is no way for Norma to ever achieve the end that she wants most desperately to achieve. Given this, we might say that rationality demands that she give up that end.¹⁶³ And indeed, I think properly valuing fame also requires her to give up

¹⁶³ We can fulfill the rational demands of a hypothetical imperative in one of two ways: by taking the means to that end, or by abandoning that end. Since Norma won’t (and perhaps can’t) take the means to achieve the end of regaining her fame, she ought, rationally speaking, to give up on her goal of making a Hollywood comeback. For more detailed discussions of the rational requirements of hypothetical imperatives see Hill, “The Hypothetical Imperative,” Philosophical Review 82 (1973): 429-50. For a different way of interpreting the
that end: if she really cares about being a Hollywood star, maybe she shouldn’t be content
with playing mother roles and accepting a spot on the celebrity C-list. Maybe it’s best that
she value fame by appreciating the celebrity she once had, and by acknowledging to herself
that she no longer has it. But this is where self-deception serves as an impenetrable barrier to
proper valuing. In her self-deceived state, Norma can’t see that she is not in fact Hollywood
royalty. She can’t, in other words, see that she hasn’t in fact accomplished the end that she
has set for herself. As such, she can never make meaningful progress towards that end:
there’s no point of trying to return to Hollywood if you think you’re already there. But she
also can’t rationally abandon her goal of returning to the silver screen. She can’t abandon a
goal that she already sees herself as having achieved. This means that Norma truly is stuck
between a rock and a hard place: her self-deception prevents her from ever properly pursuing
her goal of reclaiming her fame, and it keeps her from ever revising the content of those ends
to better reflect her values. Thus, Norma’s self-deception constitutes a failure to treat herself,
her values, her ends and ultimately herself with the seriousness that they deserve.

That, at least, is the story that an Intentionalist could tell about Norma. Now let us
consider how the Motivationalist might interpret Norma’s situation. To start, note that when
it comes to our identity-constituting ends, I have argued that self-respect requires that we
hold ourselves to especially stringent practical and epistemic standards. It’s one thing to care
about forming a justified true belief about some political matter that doesn’t really concern
you—you ought to, because as a rational agent, you should be generally committed to
forming justified true beliefs. But it’s another thing to be committed to forming a justified

standard set by the hypothetical imperative, see Mark Schroeder, “The Hypothetical Imperative?,” Australasian
true belief about a topic that interests you greatly, or that concerns one of the ends that matters most to you. When it comes to these topics, you should care about forming reasonable beliefs not simply because you’re a rational agent, but also because you’re invested in properly engaging with those topics in particular. Holding yourself to particularly stringent epistemic standards is just another way of affirming that your identity-constituting ends are worth taking seriously.

But this is exactly what Norma doesn’t do. Despite wanting badly to regain her fame, she keeps herself from paying attention to her present circumstances in the way that her end would require her to. If fame is something that she really cares about, then she should be firmly committed to seeking out genuine criticism rather than empty praise, and to watching new movies to see where she might be able to find a place for herself. But Norma instead shies away from anything that might cause her emotional pain, and hides instead in reassuring self-delusion.

Self-deception also represents a perversity of valuing in a second, higher-order sense. It is clear that Norma values knowing that she’s a successful Hollywood actress. We can see this in the fact that she demands constant reminders of her fame from those around her. She enjoys thinking about this ‘fact’ about herself, calling it to mind, and gathering additional evidence in support of its truth. And in this strange way, we can see that Norma values—in one sense at least—having true beliefs. She’s relevantly different than the washed-up actress who is content to simply daydream about what it would be like if she were still a member of the Hollywood Elite. Norma wouldn’t be content with this: she doesn’t want to imagine what it would be like if it were true that she was still famous. Rather, she wants it to be true that she is still famous. We can’t make sense of Norma’s actions unless we accept that she is, in
some perverse way, committed to self-knowledge and dissatisfied with mere fantasy. But at the same time that her actions demonstrate her to have this commitment, they also constitute a failure to honor it. And this failure, once again, is particularly tragic insofar as it is the very thing that will prevent Norma from ever actually knowing herself or discovering the truth about the world.

Section 6: Meeting the Desiderata of a Successful Account

I have demonstrated at this point that self-deception is a species of epistemic self-disrespect. Specifically, it is a failure that concerns the ends with which we most strongly identify. We are now in a position to see how my account can deal with the desiderata that I laid out at the beginning of my chapter. I’ve already talked about the sixth desiderata, which concerns the compatibility between a theory of self-deception’s epistemological structure and its moral wrongness, and so I won’t discuss it further.

Let us instead start with the first desideratum. Why is self-deception so distinct from other failures of self-knowledge? It is distinct, I think, because it typically involves a failure to properly respond to one of the ends that ought to matter most to us. Because of this, self-deception will typically involve a failure of substantial rather than trivial self-knowledge.

This is not true of self-knowledge failures generally. Often, we fail to know very trivial facts about ourselves. If I am forgetful, I might fail to remember where I put my keys or whether I enjoyed watching *Boardwalk Empire*. But whatever false beliefs I have about myself as a result of this forgetfulness do not speak to my values—I don’t have them because I’ve failed to properly engage with something I really care about, but rather because I have a bad memory.
Next, let’s consider why self-deception has the power to make a morally bad person seem morally worse. The person who is self-deceived about his moral viciousness, I suggested, is blameworthy for two reasons: he’s blameworthy first for having that character, and second for not knowing it. This double blameworthiness occurs because people who deceive themselves about how bad they really are often manage to wrong other people in addition to themselves. To go back to two characters from the beginning of this chapter, we might say that self-deceived Larry is condemnable for all the same reasons that self-aware Percy is—both are manipulative philanderers. But Larry, unlike Percy, has also failed himself. His self-deception is proof that he values, on some level, being a good guy. And it is also the very thing that prevents him from realizing that value. This explains why it is so easy to get angry at moral self-deceivers for ‘having their cake and eating it too’: they still care about being good, but their self-deception allows them to get out of doing all the work that the rest of us have to do in order to live up to our moral ideals. Percy, of course, is no moral saint: it would be much better if he actually cared about what he was like. But his failure to care about morality is still, I think, not a distinctly self-regarding failure.

These considerations help us to explain why self-deception can add moral content to otherwise non-moral situations. This is what happens in Norma’s case, when she deceives herself into believing that she is still a Hollywood starlet, and also in Penny’s, when she deceives herself about why she couldn’t complete the 10k race. Penny really cares about being fit—that’s part of the explanation of why she can’t admit to herself that her poor performance is proof that she’s out of shape. But her self-deception is also the very thing that is standing in her way of ever getting into shape. In this way, Penny fails herself insofar as
she fails to set the sorts of ends that she ought to set in order to properly engage with one of her values.

And so too are we able meet the forth desideratum, which required us to account for the range of moral and non-moral reactive attitudes that self-deception regularly elicits. We feel blame and censure towards the self-deceiver because we think, quite reasonably, that they have done wrong by themselves, and that they’ve failed to treat themselves with the self-respect they deserve. But we also feel pity and empathy. This is because, while we can think that they’ve done something morally wrong by failing themselves in this way, we can also recognize that they have, through being self-deceived, already been punished for their failure. They have, after all, been denied the ability to ever properly engage with the ends with which they most strongly identify. Depending on the circumstances, we may feel as though their self-deception is punishment enough, and that any additional censure would be inappropriately punitive.

In some situations, we may actually go further than this. We may, in fact, feel motivated to help preserve the self-deceiver’s warped worldview. This is especially common in cases where the reality from which the self-deceiver endeavors to escape is particularly horrific. Consider the classic case of the mother who is self-deceived about her terminally ill child’s chances of recovery. Her self-deception reveals what she values: that is, her child’s well-being and continued existence. And it also reveals something about how she wishes to see herself: as the mother of a child who will eventually grow into an adult. Depending on the particular form that it takes, her self-deception may actually frustrate her ability to realize that value in the limited ways in which it is possible to realize it. She may, for instance, push doctors to try experimental therapies that cause her child unnecessary pain and suffering.
instead of pursuing palliative therapies aimed at reducing pain and improving quality of life. But even if this is the case, the wretchedness of the situation might move others to help keep the mother’s self-deception alive. We are psychologically limited beings, and there is only so much mental agony that we can face head-on.

This should not be taken to imply that self-deception, even in tragic cases, counts as a morally good thing. Rather, we should appreciate that, in certain situations, the behaviors that are typically morally obligatory become supererogatory. While there might be something especially admirable about the person who remains committed to seeing himself and the world clearly in even the most tragic of circumstances, we cannot reasonably expect or demand that people respond to tragic circumstances in this way.

In this way, my account can speak to the fifth desideratum, which concerned ‘exceptional’ cases of self-deception that we do not consider condemnable. The mother’s case is one such exception, but there are other less tragic ones. Through most of this chapter, I have talked about ‘extreme’ cases of self-deception—cases involving deceptions that persist over great lengths of time and that concern serious subjects. Not all self-deception is like this. Recall the novice comedian who deceives herself about how well her first performance went. Or imagine how a job candidate might console herself at the end of a lackluster interview by musing that she didn’t really want the job anyway, or by reassuring herself that the interview didn’t go that badly.

On my account, self-deception often both reveals to others what ends we value most, while at the same time frustrating our ability to ever realize those ends. But this isn’t always the case. Sometimes, self-deception can be a means that we can take to achieve our ends. The stand-up comic and the job candidate may need to deceive themselves about their respective
performances in order to maintain the mental fortitude that they need to keep pursuing their goals.

These cases of innocuous self-deceptions are distinguishable from condemnable cases like Norma’s and Steven’s across at least four dimensions. First, as mentioned above, they are instrumentally valuable to the achievement of our ends in a way that more radical self-deception is not. If the job candidate once read a study that demonstrated that people who underestimated their chances of getting the job had a higher chance of succeeding than those who overestimated their chances, then we can imagine that she would make an effort to focus on why she won’t likely get the job before she begins the interview.

Second, these cases involve minor self-deceptions: the job candidate isn’t deceiving herself into thinking that she’s done well in every interview, but just that she’s done well in this one. Norma’s self-deception, in contrast, infects almost every aspect of her life. Third, innocuous self-deceptions are temporally limited: the teenager’s self-deception will end, one way or another, when his driving test is complete. Steven’s self-deception about the moral character of Lord Darlington persists long after Darlington is dead.

Third, innocuous self-deceptions do not require a radical misuse of our epistemic powers. The job candidate doesn’t need to work hard to reinterpret evidence about her performance; she simply spends more time thinking about the questions that she answered

---

164 The exception, of course, is if the job candidate’s self-deception is merely a symptom of a larger self-deceived belief about her competency in interviews, the attractiveness of her demeanor or so forth. In this case, it would not be apt to consider her self-deception minor. Rather, we can refer to it as a particular component of a larger self-deceived belief. And this larger belief, again, is a belief about the job candidate’s substantial self.

165 Rorty, “Belief and Self-Deception,” 400.
well than the ones she answered poorly. Stevens, in contrast, has to radically lower his epistemic standards as they relate to topics like his father, Lord Darlington, and his relationship with Miss Keaton if he is to remain self-deceived.

Finally, and perhaps most importantly, innocuous self-deceptions do not involve topics that are related to our substantial selves. The job candidate does not define herself by how well she did in the interview, but Norma does define herself by her status as a famous actress. And so, while the job candidate’s self-deception involves rather trivial beliefs about herself, Norma’s self-deception causes her to lose sight about the very parts of herself that self-love and self-respect demand she try to know.
CONCLUSION: WHAT FLORENCE SHOULD HAVE KNOWN

I have argued that a non-instrumental commitment to knowing oneself is not a sign of morally condemnatory narcissism, but rather of appropriate self-love. A willingness to take seriously that commitment, by holding oneself accountable to appropriately stringent epistemic standards, further, is part of what it takes to respect oneself as an epistemically rational agent.

In the process of establishing these claims, I have shown that self-respect does not frustrate self-love, as some might think, but actually compels us to love ourselves. This is because self-respect requires both that we exercise a reasonable partiality towards ourselves and the ends that we set, and also that we strive to act in ways that make us loveable to ourselves.

Finally, I have explained why we need not appeal to consequential considerations to explain our antipathy to self-deception. Rather, self-deception is in and of itself an undesirable state insofar as it constitutes an especially objectionable form of epistemic disrespect. The self-deceived agent does not simply fail to adequately scrutinize her beliefs. Rather, she fails to adequately scrutinize the very beliefs that are most deserving of her epistemic attention, viz. the beliefs that relate to the ends, values, and commitments with which she herself most intimately identifies. Self-deception thus often serves to reveal to others which of our ends we value most, while at the same time frustrating our ability to ever properly engage with those ends.
Putting the pieces of this puzzle together, we can now explain what reason Florence would have had to learn about her lack of musicality. As I made clear in the Introduction, this reason could not be an instrumental one: Florence would not have been able to use this newfound self-knowledge to improve upon herself or her life, and would have no doubt found it to be painful. And yet, I think many of us still feel pulled towards the idea that Florence had *some* reason to want to know the truth about her singing abilities, and that she stood to gain something by acquiring this self-knowledge.

The reason that Florence had to know herself is the same sort of reason that we have to discover unconvertible truths about our loved ones, even when those truths are likely to harm the relationship we share with them and cause us emotional pain. Our love for ourselves gives us reasons to know ourselves that don’t bottom out in considerations of well-being, and our self-respect pushes us to act on those reasons. If Florence had wanted to discover this unpleasant truth about her musicality, then, we could say that she respected and loved herself enough to be committed to seeing herself clearly.

This does not mean, however, that Florence’s self-discovery, as a matter of historical fact, was a net good for her. Nor does it imply that it was a good thing that the self-discovery occurred in the precise way that it did, namely after reading a review of her Carnegie Hall performance that her romantic partner had tried to keep hidden from her. Rather, we can say that it was a shame that Florence’s self-discovery came so late, insofar as St. Clair had for years struggled to deny Florence any opportunity to see herself clearly. And insofar as Florence herself also had a role to play in the ruse, we can say that it was a pity that she didn’t take one of her most identity-defining ends, and thus herself, more seriously. Finally, even if Florence was made happier by living a lie, we can still contend that she lost out on
something valuable. Specifically, she lost out on any opportunity she may have had to know
the person with whom she shared her most intimate relationship.
BIBLIOGRAPHY


