

Routledge Studies in Contemporary Philosophy

Updating the Interpretive Turn
New Arguments in Hermeneutics
Edited by Michiel Meijer

Conservatism and Grace
The Conservative Case for Religion by Establishment
Sebastian Morello

The Ethics of Interpretation
From Charity as a Principle to Love as a Hermeneutic Imperative
Pol Vandeveldde

The Nature and Practice of Trust
Marc A. Cohen

A Plea for Plausibility
Toward a Comparative Decision Theory
John R. Welch

Living with the Dead
On Death, the Dead, and Immortality
J. Jeremy Wisniewski

Free Will's Value
Criminal Justice, Pride, and Love
John Lemos

Emotional Self-Knowledge
Edited by Alba Montes Sánchez and Alessandro Salice

For more information about this series, please visit: <https://www.routledge.com/Routledge-Studies-in-Contemporary-Philosophy/book-series/SE0720>

Emotional Self-Knowledge

Edited by
Alba Montes Sánchez
and Alessandro Salice

 **Routledge**
Taylor & Francis Group
NEW YORK AND LONDON



2 Alienated Emotions and Self-Knowledge

Krista K. Thomason

I might be afraid of mice. I say “might be” because I have conflicting evidence. On the one hand, when I was a child, I had a mouse as a pet (her name was Rocky). I would take her out of her cedar chip-filled aquarium and happily hold her. I loved her like any other pet and mourned her when she died. On the other hand, there is what I call “the kitchen incident.” As an adult, in the first house I owned, I once heard odd noises coming from my kitchen at night. I gingerly walked down the hall from my bedroom and flipped on the kitchen light to look around. A mouse scrambled across my bare foot. I screamed so loud that I woke up my husband. I was so paralyzed with fear that he had to take my hand and physically walk me back to the bedroom because I refused to leave the safety of the kitchen mat next to the sink.

Since the kitchen incident, I do not how to answer the question “Are you afraid of mice?” Prior to the kitchen incident, I would have answered, obviously no. How could someone who had a pet mouse be afraid of mice? I would have bristled at the stereotype: the familiar cartoonish image of a woman standing on a chair in terror of a little harmless mouse. Back then, if you had asked me how I would have reacted to a mouse in the kitchen, I might have said that I would be startled, but I never would have predicted my actual reaction. My fear was completely surprising to me. In the moment, I simply found myself paralyzed on the kitchen mat screaming my head off and unable to take a step without the help of another person.

My goal for this chapter is not to determine whether I am afraid of mice. I want to explore what I will call the ethics of emotional self-knowledge. I am particularly interested in how we relate to what I will call *alienated emotions*: emotional experiences that are unusual, surprising, or even disturbing.¹ Alienated emotions are those that make us feel “momentarily or robustly alienated...from one’s own ‘self’” (Szanto 2017, 264). What, if anything, do our alienated emotions tell us about who we are? Before I proceed, I need to make a few clarifications. First, I am thinking of self-knowledge in a broadly moral sense and in terms of what Cassam calls “substantial self-knowledge” (2015, 10). I will assume that trying to gain self-knowledge is a good or worthwhile thing to do, but I do not

want to cash it out in narrow moral terms, such as duty, obligation, or permissibility. Instead, I will characterize it as a practical project: something human beings do as they figure out how to live meaningfully in the world. As such, I will set aside some of the more foundational questions from philosophy of mind and epistemology about, for example, privileged access or self-ascriptions of mental states. Second, I do not want to assume at the outset any particular view about what self-knowledge looks like, though I am primarily concerned with our self-concept. Because of that, my affinities with normative narrative selfhood will no doubt come through.² I do not wish to stake out a position about what kind of narrativist I am (if I am one), and, in particular, I do not intend to make any strong claims about whether self-understanding is essentially narrative in structure. I will return to some of these issues later in the chapter. Finally, I will use several terms interchangeably: self-knowledge, self-discovery, self-understanding, and self-examination will all pick out basically the same process or the same goal.

2.1 Emotions Tell Us Things about Ourselves

Let me start with a claim that I take to be relatively uncontroversial, namely that emotions do indeed tell us things about ourselves. There are at least two ways they might do so. First, emotions reveal what we care about or what we judge.³ Michael Stocker argues that caring about something is essentially an affective activity—what it means for me to care about baseball is to be sad when I do not get to watch it, thrilled by an exciting game, or angry when my team makes stupid roster moves.⁴ Robert Solomon and Angela Smith argue that emotions are tied to our judgments.⁵ The fact that I feel contempt for my neighbor means that I have judged her to be beneath me or the fact that I scream my head off when the mouse runs across my foot means I judge it to be scary.

Second, emotions reveal our character traits. This claim has a long history and is featured in Confucius, Aristotle, and Hume. Particularly for Confucius and Aristotle, the person with the right sort of character will feel the emotions that are appropriate to that character (e.g., the courageous person will feel the right amount of fear in the right way and at the right time).⁶ We take someone's emotional response to be evidence that they have developed a certain trait (either a vice or a virtue). We can see this via an analogy between emotions and actions. As Hume argues, we blame people for their "criminal actions" because they are "proofs of criminal passions or principles in the mind" (2007, 2.3.2, 264–285). We infer, in other words, that someone's actions are produced by their character. The same can be true of our emotional responses. We can infer that someone who is prone to fear or who feels fear at, say, a small harmless rodent, is probably a coward. Although Hume talks in terms of blame, the same thing can be true of positive emotions. Someone who

weeps for another's suffering is likely to be praised as a sympathetic or compassionate person.

Of course, both of these connections immediately get complicated. First of all, how precisely are the connections supposed to work? Why and how does an emotion follow from my judgment? How exactly does an emotional response "flow" from a character trait?⁷ Second, the connections themselves are tenuous. Emotions can be recalcitrant—how we feel can conflict with what we judge.⁸ I might sincerely think that mice are not dangerous and yet, lo and behold, I am afraid when I see one (I will come back to the issue of the alleged irrationality of recalcitrant emotions in a moment). In the same way that we can act out of character, we can also feel out of character. A person with a normally generous, sunny disposition can feel malicious glee or a sudden sullenness, maybe even to her own surprise.

So, we end up with the following problem: sometimes emotions tell us things about ourselves and sometimes they do not. How do we know which is which with any particular emotion? Is the kitchen incident telling me that I am afraid of mice or not? When we confront conflicting pieces of evidence in our quests for self-knowledge, it provides us an opportunity for self-interpretation.⁹ But self-interpretation is a fraught task, especially when it comes to emotions. First, there is the problem of opacity: we are not transparent to ourselves.¹⁰ I do not always know what I feel, why I feel it, or at the very least I have trouble articulating it. Second, there is the problem of confabulation. Psychological research purports to show that people will invent plausible, but false, explanations for their own behavior.¹¹ We tend to create reasons for our actions that may not have been the reasons why we in fact acted. We do the same with our emotions: suppose I am feeling sad for no particular reason. I might try to find reasons for my feelings, but in doing so I actually invent reasons (e.g., I decide that I do not like the paint color in my living room even though it has never bothered me before). Finally, and relatedly, there is the problem of self-deception.¹² My emotions may indeed tell me things about myself, but they may tell me things I do not like. If I do not want to face up to these unpleasant facts about me, I may dismiss the feelings as an aberration. If I am motivated to see myself in a certain way, I might ignore or downplay feelings that present me differently.

If self-knowledge is a good thing and we ought to try to achieve it, I take it we should try to avoid or address these pitfalls, if we can. Just as there are better and worse ways to deliberate, there are better and worse ways to engage in self-interpretation. How do we self-interpret well when it comes to our emotions? In what follows, I want to try to make some headway into this question by examining cases of alienated emotions, like the kitchen incident. I want to first identify some tempting responses we might have to our alienated emotions and show why those responses might be problematic. I will then try to identify what I will call guiding values of self-interpretation.

2.2 Just a Reaction

Here is, I think, the first tempting response: my fear in the kitchen is an aberration because it's just a reaction in the moment. It was late at night, I was already on high alert because I was hearing strange noises in the house, and out of nowhere a furry, fast-moving thing runs across my bare foot. I can imagine many people who are otherwise not afraid of mice might react the same way. I think fear lends itself to the "just a reaction" conclusion partly because it's one of the so-called basic emotions.¹³ The idea that there are emotional building blocks isn't new—both Descartes and Spinoza held something like a basic emotion thesis. The basic emotion thesis is the idea that all human beings are born with a set of emotions. This claim is supported by studies in emotion recognition: show people from a wide array of backgrounds and regions a bunch of photos of human faces with different emotional expressions and ask them to identify the emotion in the photos.¹⁴ At least according to some studies, subjects are successful at this emotion identification task. Since we can identify emotions across cultures, basic emotions are supposed to be "hardwired" into human psychobiology.¹⁵ This claim is then supported by evolutionary psychology, which tries to determine what evolutionary function the basic emotions are meant to play. Fear, for instance, causes your blood to flow more freely to your large muscle groups.¹⁶ More blood flow makes it easier to run, so presumably feelings of fear evolved to help us flee from danger. If all this provides an accurate description of fear, then perhaps my fear of the mouse in the kitchen incident means very little—it was a psychobiological reaction, like flight or fight, nothing more.

When it comes to emotions and self-knowledge, I think this answer is unhelpful, both in general and also in my particular case. To start with, although this is not so central to my arguments there, it is important to note that the research on basic emotions is not unassailable. The psychologist Lisa Feldman Barrett, for example, argues against the basic emotion thesis (2017, Chapter 3). Her research raises questions about the results of the facial recognition studies. In some of the original studies, subjects were asked to choose the emotion depicted in the photo by selecting from a list. Barrett ran the experiment without the list; her researchers just asked people "what emotion is this person feeling?" Without the list, subjects had a much more difficult time identifying the emotion (2017, 44–46). Barrett thus doubts that there are basic emotions and that they are "passed down" from our earlier ancestors (2017, 157–174).

It is not central to my argument whether there are or are not basic emotions, but it is important to think about how classifications like these might influence how we relate to our emotions. One of the downsides of classifying an emotion as basic is that it sometimes ends up getting cast as simplistic. Fear becomes equated to the flight-or-fight mechanism or the startle reflex. Of course, the basic emotion thesis is meant to make

emotions reductive to a certain extent: the point is to identify the simpler parts that make up the more complex emotional structures. For that reason, it focuses on more simplistic cases of fear rather than, say, thinking about the fear of death, the fear that your loved ones secretly resent you, or the fear that you are becoming exactly like your mother. Because of this, we tend to look at unexpected fear responses as not particularly deep and therefore easier to dismiss as aberrations.

A second reason to doubt the "just a reaction" interpretation is that emotions help us discover things about ourselves that we do not already know. An unpredictable emotional response may nevertheless be a genuine one. You might learn that you are afraid of heights only after you have decided to hike up a mountain or ride Ferris wheel. You might discover that you harbor romantic feelings for someone only upon hearing that they have a date planned with someone else. If I had never encountered a mouse prior to the kitchen incident, I might easily conclude from my reaction that I am afraid of them even if I did not realize it beforehand. Of course, one emotional reaction all by itself might not be enough evidence to draw wider conclusions. You might need to try hiking a few more times to know if you are afraid of heights. Still, if your fear is serious the first time—if you are terror-stricken on the summit—the claim that it is "just a reaction" might be self-deception. In the kitchen incident, I was (to my utter surprise) screaming uncontrollably and quite literally paralyzed with fear. I was not merely startled; but I was also in a full-blown panic. The severity of my fear makes it not only shocking, but also much harder to dismiss. Even when our emotional reactions are surprising or unpredictable, it does not mean they are mistaken.

2.3 Recalcitrant Fear

Another tempting way to think about the kitchen incident is to say my fear was recalcitrant. An emotion is recalcitrant because it is "at odds with a decisive better judgment of the subject" (Mele 1989, 279) or "persists despite the agent's conflicting judgment or belief" (Döring 2015, 381) or "exists despite the agent's making a judgment that is in tension with it" (D'Arms and Jacobson 2003, 129). Most of the literature on recalcitrant emotions uses fear as an example: Patricia Greenspan's landmark case is about someone who had a terrifying incident with one dog and then develops a general fear of dogs (1988, 17–18). This person then meets Fido, the harmless old arthritic dog. The person knows that Fido is not dangerous and yet fears him anyway. Recalcitrant emotions are the kind that we know we have no reason for, that we ourselves sometimes claim are irrational, and that persist despite our other beliefs and judgments.

In my example, the case for recalcitrance would go like this: I went through my whole life having positive attitudes toward mice. Given this,

and given that I know mice are not dangerous, my considered belief or judgment is that I am not afraid of them. The fact that I was panic-stricken in the kitchen incident could have a number of different explanations, but it did not represent my real, authentic, or considered attitude toward mice. So, in my quest for self-knowledge, I can safely disregard it.

Classifying an emotion as recalcitrant is harder than it looks. Let me borrow from Smith's rational relations view to illustrate.¹⁷ Smith's view provides a way of explaining how moral agents can be responsible for their attitudes that does not rely on tracing those attitudes back to some sort of choice, endorsement, or control.¹⁸ We can and do have attitudes that we do not choose or endorse, but on Smith's view, we can still be answerable for them and open to evaluation because of them. What makes our emotions open to evaluation is that they reflect our values in the right ways—they "bear rational relations to our evaluative judgments and commitments" (2005, 260). Contrast my fear in the kitchen incident to the startle reaction I might have when I unexpectedly run into someone coming around the corner. I cannot be called to account for my startle reaction because it is not attached to any values I have. No one will look at my startle reaction and say, "Wow, you must think people walking around corners are dangerous." A startle is not a reflection of judgment or a commitment.

As Smith points out, your judgments could be deeply mistaken and yet your emotions still be connected to them in the right ways (2005, 253–254). Suppose I am afraid of mice because I falsely judge that they are venomous. My fear of them is indeed a reflection of my judgment and the fact that venomous things are dangerous even makes my fear reasonable under that description. I am just wrong about the facts regarding which animals are venomous. If my judgment gets corrected and yet my fear persists, then my fear would be a classic case of recalcitrance because it would properly conflict with my considered judgment. When this occurs, Smith thinks we are faced with "an interpretative difficulty" because we cannot trace a clear route from the emotion to the judgment (2005, 255).

My fear in the kitchen incident is not like the startle reaction in that it might actually reflect my judgments. Why suppose this? I think the intensity of my response is one reason. When I run into someone coming around the corner, I might jump and even scream a bit, but it is over quickly. I might even laugh about it immediately afterward. The intensity of my fear in the kitchen incident has not been so easy to get over or laugh off. My views about mice have been unsettled ever since. Because my reaction was so shocking to me, I have since wondered what would happen if I came in close contact with another mouse. I no longer trust myself around them and I am still amassing conflicting evidence. I have been to pet stores where they are in aquariums and had no fear response. I have seen them outside while walking my dog and have not been gripped with terror, but I was uncomfortable. I could not help trying to

make sure I kept sight of them and worrying that they would suddenly scurry back across my path. I read a news article about a man who fell through the cracked sidewalk in New York City into a pile of rats and I nearly break into a cold sweat every time I think about it. On television, I have seen videos of lots of field mice running around in Australia and had to turn my head away from the screen. So, although the terror I had in the kitchen incident has not reoccurred, it has forced me to pay closer attention to and re-evaluate my other responses. Although no one else has judged me for my fear (partly because I have not told that many people), I absolutely judge myself or at the very least I see myself as open to judgment. I can imagine people asking me the very same questions I have asked myself: are you *that* afraid of mice? Did you think it was going to hurt you? Why did you react that way?

Now comes the crux of the problem. On the rational relations view, we are answerable for our emotions when they reflect what we judge, yet in my own case, I cannot figure out where my fear came from. Up until the kitchen incident, I would have identified nothing in my judgments or commitments that would have predicted my response. I am an animal lover, I am not prone to terror over other things like bugs or spiders, and I used to have a mouse as a pet. My emotion conflicts with all of this, but does that mean that my fear is floating free from the rest of my psychology? Or is it that I do not know myself as well as I thought?

When it comes to self-knowledge, the mere fact that an emotion conflicts with our considered judgments does not all by itself count as a reason to disregard it.¹⁹ First, as Smith points out, our conflicted attitudes are not "alien forces" that operate within us (2004, 399). When we struggle with feelings that we wish we did not have, we nevertheless recognize them as ours. Who I am as a person is comprised of things that I do not necessarily choose: where I grew up, the activities that I enjoy, and the people I love are not always hand-selected by me only after careful consideration. We do not choose our families, but (most of the time) we love them anyway. I stumbled into philosophy almost by accident and fell madly in love with it. We don't decide what feelings to have because emotions simply do not work that way, but that does not make them not authentically ours. The sphere of my identity is bigger than my willing and deliberating.²⁰

One of the common conclusions we draw about recalcitrant emotions is that they are irrational. If that is right, so this reasoning goes, we do not need to care much what they might tell us. I have argued elsewhere that it is hard to determine how recalcitrant emotions are irrational.²¹ The case in favor of irrationality for recalcitrant emotions relies heavily on the notion of emotions having formal objects.²² Fear is supposed to track or perceive "the dangerous." But fear that looks irrational might be tracking something that is outside the scope of an emotion's usual formal object. Take the fear of spiders as an example. What if someone fears spiders not because they are dangerous, but because they are

creepy-crawly? Again, we tend to construe fear's formal object in a narrow way, as though "the dangerous" can only include things that can cause one serious physical harm. But we are also afraid of the disgusting, the eerie, the creepy, or the uncanny.²³ You see a spider and you watch its little hairy legs creep across the floor; you imagine that if you get too close, it might suddenly scurry across your foot or up your leg. Notice that the fear of the creepy-crawly is not alleviated by knowing that the spider is not venomous—creepy-crawlies do not have to be venomous to be scary. Fears that look irrational might not be once we realize that they are attuned to atypical objects. For the purposes of self-knowledge, we would be better served not to apply the recalcitrant label too quickly.

Even if it is true that recalcitrant emotions are irrational, why should we merely disregard them for the purposes of self-knowledge? An irrational instance of an emotion is still an instance of that emotion. As Pugmire puts it, "Irrational guilt is not pseudo-guilt" (1998, 123). Irrational emotions are not fake, feigned, or false, even if we have no clear reason to feel them. Objectless emotions illustrate this: we can be sad or happy without being able to explain exactly why, but objectless sadness and happiness are still classified as sadness and happiness.²⁴ Objectless emotions do not seem to strike people as particularly irrational, unless they persist for a long time or are very intense. Even if I judge my own emotion to be irrational, it could still tell me something about myself. Suppose I feel guilty for turning down yet another service request, even though I am not interested in the project and I have too much to do already. My irrational guilt might be revealing all sorts of things about me—that I am a people-pleaser, that I agree to things for the wrong reasons, or that I am struggling with wanting to do too much.

Additionally, the fact that an emotion conflicts with my judgments does not mean my considered judgments are correct. As I mentioned earlier, our emotions can alert us to things about ourselves that we do not already know and sometimes they know better than we do. Let me illustrate with a case from Nomy Arpaly (2000, 496–498). Sam is a college student worried about doing well on final exams. He decides that he should restrict his social life to the absolute minimum so that he can study as much as possible. In his deliberations, Sam fails to realize that when he does not maintain a balanced social life, his academic performance suffers because he becomes depressed. His considered judgment about what is best for him is that he ought to become a hermit during finals, but Arpaly argues that Sam would be more rational or make a better overall decision if he acted *against* his considered judgment.

Arpaly's case demonstrates that we are imperfect reasoners. Deliberation is a messier thing than the philosophical literature sometimes makes it seem. We might think that we have come to a good conclusion about our considered judgments and values, but we can be wrong about them. We can also be self-deceived or unduly influenced by other

people's expectations about what is best. We tell ourselves stories about what we think we want and who we think we are. Emotions that conflict with considered judgments might be giving us a more holistic picture of ourselves and our situations. Suppose Sam is irritable and depressed during his isolation. He might attribute his feelings to the stress of finals. He might chastise himself for being weak-willed when he longs wistfully to join his friends for dinner. But his emotions—his anger, sadness, and longing—are pointing to a truth of his situation that he is unable to realize or see. This sort of thing happens all the time. If you are crying to and from work every day, it might be because you hate your job. If you dread seeing your family, your relationship to them might not be as healthy as you think. Of course, not every conflicting emotion is like this—my boredom and lack of focus does not necessarily mean that I hate what I'm working on. But the mere fact that a feeling goes against our considered judgment does not mean the feeling is the mistaken one.

So far, I have tried to make the case that we should not just ignore or dismiss emotional reactions that seem like aberrations. They may not be just reactions and they may not be recalcitrant. Even if they are irrational, that does not disqualify them from the project of self-knowledge. Emotional reactions that are shocking, disturbing, or surprising might still reveal something about ourselves. The troubling word here is "might." How do we know when an emotion is telling and when it is not? I'm not sanguine about the possibility of actually answering the question I just posed, but I think there are better and worse ways of trying to answer it. In this final section, I want to try to sketch some general principles for good emotional self-discovery. When we are trying to understand ourselves, how *should* we reflect on our emotional experiences?

2.4 What Does Good Emotional Self-Knowledge Look Like?

The conclusion I want to draw from what I have argued so far is that there is no reason to prioritize our considered judgments over our emotions in cases of emotional self-alienation. I think we have a tendency to assume that a surprising emotion—particularly when it is a strong one—is the problem. My main goal in this chapter has been to argue against this tendency. There is no reason to think that a surprising emotion is a fluke that tells me nothing about myself, and there is no reason to assume that my considered judgments are ironclad. Although alienated emotions are shocking or surprising, it is not the mere fact that they are surprising that involves them in self-knowledge. I might be excited to go to a museum and yet nevertheless the experience leaves me cold. This is not what I expected, but it need not threaten my sense of myself as someone who loves museums (unless it happens over and over again). What stands out about the kitchen incident is that my fear was contrary to a relatively clear and stable self-image that I had good reason to

believe was authentic. Up until the incident, I had no reason to imagine that I would be terror-stricken at the sight of a mouse in my kitchen and no reason to think of myself as a person who might be afraid of mice. My feelings of fear have made me doubt some of the self-knowledge or self-clarity that I thought I had obtained.

If we accept that emotions can reveal things to us about ourselves and that they might be more revealing than our considered judgments, then alienated emotions could be telling the truth about who we are. This is what can make them upsetting or disturbing. When people develop romantic feelings for someone other than their partners, when they respond with a sudden burst of rage or cruelty, when they react with disgust toward someone they thought they cared about, these are all possible examples of emotional alienation. Alienated emotions are the ones that haunt us—the ones that make us doubt that we actually know ourselves as well as we thought. These moments of self-doubt should not be dismissed or ignored just because it is a strong emotional response that occasions them. I think we ought to take experiences of emotional self-alienation seriously when we are trying to gain self-knowledge. When you have an emotion that is shocking or surprising to you, you ought to take it as a call to engage in self-examination rather than dismiss the feeling as an aberration. Sometimes feelings are just feelings, but sometimes they are more. There is no a priori way to know which is which unless we engage in self-interpretation.

My claim here is essentially a normative one. That is, we can be better and worse at our quest for self-knowledge, and if we want to do it well, we ought not dismiss our alienated emotions. Alienated emotions help highlight the challenge we face in engaging in self-interpretation well. Self-interpretation about alienated emotions can be particularly fraught because of the way the emotions conflict with our self-image.²⁵ By “self-image” here, I just mean the way that I understand myself. A self-image can be more and less stable, and it can be revised over time. I can also get my self-image wrong: I can be incorrect or deceived about who I am. Alienated emotions surprise or confuse us because they seem to conflict with who we think we are.²⁶ Prior to the kitchen incident, I understood myself as someone who was fond of (and decidedly not scared of) mice. My fear response thus felt alienated. Not only did I have the self-image of someone who was not afraid of mice, but I also had become *invested* in that self-image. We can invest in our self-images for a number of different reasons—some flattering and some not-so-flattering. One of the possible reasons I was invested in my mouse-loving self-image is because I could safely inoculate myself against the sexist stereotype of women who are afraid of mice. Those of us who are members of marginalized groups can be overly sensitive to stereotypes—we fear falling into them and often try to cultivate the opposite in ourselves. I might have

taken a secret pride in my sense of myself as “not one of those women,” and I might have only been nascently aware I felt this way.

It is easy to see how our investment in our self-image can stem from motivated reasoning, but such investment need not be so nefarious. Coming to clarity about who I am is intimately tied to my ability to be a person in the world. Knowing what I value, what I desire, and what I hope to accomplish in my life is inseparable from having some sort of understanding of myself. Identity crises are painful in part because they make it hard to go forward with our life—not knowing who you are can mean that you no longer know what matters to you. There is nothing self-deceptive about wanting to have some stable sense of who you are, but that desire can lead us to dismiss or ignore emotional responses that conflict with our self-image.

Given the fact that alienated emotions conflict with our self-image, we are especially prone to bad faith or self-deception in interpreting them. How might we avoid these pitfalls? It would be helpful to have a sense of what values should be guiding our search for self-knowledge. Although I cannot offer a full account of good self-interpretation here, I want to offer three possibilities for values or principles that might serve as guides: accuracy, consistency, and honesty. The first two are, I think, problematic while the last one gets closer to the mark.

Let me start with *accuracy*. When I am thinking about my fear in the kitchen incident and weighing it against my self-conception as a person who likes mice, I do not think I am doing this just so I can settle the question “am I afraid of mice?” Of course, I *do* want to know whether I am afraid of mice. Some of my reasons are practical—if I am afraid of mice, it would be good to know so that, for instance, I do not suddenly decide to switch careers and become an exterminator. Some of my reasons are purely for the sake of knowing myself. I do not like the idea that my fears are mysterious to me and I want to have clarity about the sort of person I am. But knowing myself is not identical to amassing a collection of true facts about me, like an encyclopedia entry. Self-knowledge is also not a particular instance of the overall goal of having true beliefs. It is not as though I want to have true beliefs about the world, and since I am part of the world, I also want to have true beliefs about me. First, self-knowledge is practical rather than theoretical: knowing myself is intimately tied to how I live and what I care about. Second, the thing that I am trying to know is not a stable object. Unlike facts about tigers, what I learn about myself changes how I relate to myself, which is to say, it changes me. Self-knowledge is not best understood as merely a subset of world-knowledge.²⁷

Another value that I think is a false start is *coherence*. Here is where I think some of the stronger versions of narrative selfhood get into trouble. I have in mind positions that Dan Hutto refers to as “strong narrativist” (2016, 26). Purportedly, strong narrativists hold some combination

of the following claims: that the self is essentially narrative in structure (that having or making a self is to tell a story) and that we ought to (construed broadly) engage in story-telling so that we have or maintain a sense of self. The first claim is ontological and the second is normative. I cannot argue against these claims here, but I think strong versions of both are implausible.²⁸ I would deny, as Dan Zahavi writes, that “*the self is a narratively constructed entity and that every access to self and other are mediated by narratives*” (2007, 184). The primary worry I want to raise here is against the normative claim—that we *should* try to tell a story about who we are and aim at coherence.

There are, I think, two ways to understand coherence, one of which I do not want to argue against. As I mentioned earlier, I agree that human beings are sense-making creatures who try to understand themselves as they live in the world, so to this extent, we are guided by some notion of coherence. The stronger understanding of coherence is that we ought to (as an ethical matter) aim at a kind of “narrative unity” of ourselves and our lives (Rudd 2009, 67). This form of coherence is not only more substantive, I take it, but it also has no built-in defense against self-deception. Let me be clear that my objection here is not to narrativity in general: I do not mean to suggest that narrativity is inherently unreliable, motivated, or selective.²⁹ Rather, the trouble is that the imperative that you organize your life into a narrative does not automatically mean that you will construct that narrative well. The idea that I ought to be able to tell a unified or coherent story about who I am does not prevent me from telling an inaccurate, self-aggrandizing story. A self-deceptive story is still a coherent story. In order to address this worry, narrativists have argued that there is some robust relationship between self-narrative and the good.³⁰ This argument is meant to ward off the possibility of someone living a unified narrative that is also a bad one—a supervillain might have narrative unity of self and yet be vicious.³¹

But my concern is different: if we are supposed to aim at coherence, there is no built-in guidance for what to do with parts of ourselves that contradict that narrative. Until the kitchen incident, I had a narrative unity of myself as someone who was not squeamish about rodents. My terror in the kitchen disrupts that narrative. If I am aiming at coherence, what is to stop me from simply dismissing that response as an aberration and jettisoning it from my story? This is especially tempting in my case because I had perfectly good non-deceptive reasons for constructing my narrative as I did and I do not want to think of myself as someone who is squeamish about rodents. Taking coherence as my guide would, I think, push me toward explaining away my shocking or surprising reactions rather than taking them seriously.

For this reason, I want to propose that one of our guiding values in the quest for self-knowledge ought to be *honesty*.³² We are used to thinking of honesty in the context of speech and the context of virtue, but less so as it

is related to self-knowledge. Honesty, on my view, is a way to undertake the project of self-knowledge.³³ It is a commitment we make to try to see ourselves as we are and not how we want ourselves to be. Honesty requires an openness to all the parts of ourselves, including the not-so-flattering, quirky, dark, and obscure parts. Being honest means that we accept the limitations of our self-knowledge without giving up on the project.

Honesty as a guiding value has advantages over both accuracy and coherence. First, honesty is compatible with self-opacity. We can be confused, deceived, and surprised by ourselves, which can make accurate self-knowledge hard to come by. Taylor and Moran point out that there is a special relationship between our emotions and our sense of, awareness of, and evaluations of them.³⁴ The fact that I realize that I am angry has the potential to change my feelings of anger—it can weaken them, strengthen them, or cast them in a new light. As I mentioned earlier, my terror in the kitchen incident has made me re-evaluate my sense of myself, my other fears, and my memories of past events. Because of the way our inner mental life is layered on top of itself, there may be depths of our psychology that we cannot plumb. Honesty simply requires that we acknowledge this, which means that we cannot simply dismiss without further reflection responses that seem shocking or surprising to us. We also have to be honest about the obstacles to self-knowledge that we put in our own way. For example, because I am sensitive to the feminine stereotype of women who are scared of mice, I may cling a bit too hard to my self-image as someone who breaks that stereotype. I may be loath to question it or let it go, even when my emotional responses go against it.

Honesty also works better than coherence. It allows space for emotional conflict and ambivalence. We occupy, as Amelie Rorty puts it, “a wide variety of modes of life,” not all of which move seamlessly together (2010, 427). Too much focus on coherence and resolving conflicts may make us insensitive to bits of self-knowledge that present themselves only when things are not in harmony. Feeling conflicted about something is sometimes a way of acknowledging that we are faced with two important and competing values. Resolving our feelings too quickly might require that we falsely downplay one of the values.³⁵ Honesty also helps resolve some of the worries about self-deception in coherence. Using honesty to guide self-knowledge would make us less prone to rationalizing away our alienated emotions. In my own case, the more honest answer to the question “am I afraid of mice?” might be “I feel conflicted about them.” I might feel some practical pressure to resolve this conflict so that I can, for example, avoid situations where there will be mice. But, in terms of self-knowledge, it may be more honest for me to simply accept my ambivalent feelings. Honesty also does not require that we ensure our selfhood story is always unified. Instead, we ought to see our stories as open-ended and possibly changing direction. Because we are at times opaque to ourselves, there is an ineliminable element of discovery in the

project of self-knowledge. If we are honest, we will take those discoveries as they come and explore them without worrying too much about how they fit into the story we've already written.

Am I afraid of mice? I do not know, but settling the question is less important to me than being honest about it. Alienated emotions can still tell us about ourselves even when we do not recognize ourselves in them. Self-knowledge involves, as Solomon puts it, an "ongoing emotional-reflective process" where our emotional responses and our sense of ourselves develop and change together (2007, 265).

Notes

1. I constructed this term by rephrasing Szanto's term "emotional self-alienation" (2017, 262). His use of the term is different from the way I am using it in this chapter.
2. For an overview of this literature, see Crone (2020).
3. For a small sample of literature that argues this, see Stocker (1996), Smith (2005), Solomon (1993, 2007), Baier (2010), Helm (2001) and his chapter in this volume.
4. Stocker (1996, 56–57).
5. Solomon (1993, 125–127) and Smith (2005, 249–250).
6. Confucius (2003) and Aristotle (2002). Compare with the classical Indian views from MacKenzie's chapter in this volume.
7. I take the term "flow" from Wolf 2015, 357.
8. For just a sample of that literature, see Greenspan (1988), Mele (1989), D'Arms and Jacobson (2003), Rääkkä (2005), Brady (2009), Benbaji (2013), Döring (2015), and Helm (2015).
9. I am borrowing this term from Taylor (1985), Solomon (2007), and Smith (2005).
10. The literature I am drawing on here is primarily from Kant scholarship on moral self-knowledge, known as the "Opacity Thesis." See, for example, Ware (2009).
11. Confabulation is primarily discussed in literature having to do with straightforward ascriptions of mental states. See Scaife (2014) for a helpful overview.
12. The literature on self-deception is enormous. See Vendrell Ferran's chapter in this volume.
13. Solomon has a helpful chapter on this (2003, Chapter 8).
14. These experiments are described in Barrett (2017, 4–8).
15. For a review of some of the seminal papers in this literature and critiques of them, see Solomon (2003, 122–124).
16. This example comes from Goleman (1995, 6).
17. That view is spelled out in a number of papers, see Smith (2004, 2005, 2008).
18. For critiques of her attempt, see Fischer and Tognazzi (2009) and Shoemaker (2011).
19. This section is drawing on work I have done in another paper, see Thomason (2022).
20. I have argued elsewhere for this, see Thomason (2018, 87–94).
21. Thomason (2022).
22. For discussions of fit, correctness conditions, and formal objects, see D'Arms and Jacobson (2000).

23. Windsor (2019) argues for a claim like this about the uncanny, although he maintains that we find the uncanny threatening in some way.
24. See Lamb (1987) and Price (2006) for helpful discussions. Both Lamb and Price distinguish objectless emotions from moods. Although moods are also objectless, there can also be, as Lamb puts it, "emotion-instances" that are objectless (1987, 108). For example, I can not only be in an anxious mood, but I can also experience a sudden nameless dread.
25. See Bortolan in this volume.
26. This is true even when the emotion is authentic and our self-image is deceived. Suppose, for example, that I am convinced I am not the sort of person who dates musicians, and yet I find myself attracted to a guitar player. My attraction can be authentic, but if it conflicts with how I see myself, I will *experience* it as alienated.
27. As Moran points out, even when we are trying to identify what we feel (and so, in some sense "get the facts"), taking a purely theoretical interest in your emotions alienates you from them (1988, 141–142).
28. The classic arguments against narrativity obviously come from Strawson (2004).
29. I'm drawing on Strawson's discussion of "revision" here (2004, 442–445).
30. Taylor (1989, 51–52) and Rudd (2009, 69).
31. Rudd interprets the concern this way (2009, 69–71).
32. I don't mean to suggest that honesty should be our only guiding value.
33. My thinking about honesty is drawing on the work of Montaigne and Nietzsche. For a good reconstruction of Nietzsche on honesty, see Harper (2015). For secondary literature on Montaigne, see Shklar (1984, chap 5) and Bakewell (2011).
34. Moran (1988) and Taylor (1985). They cash this relationship out in different ways.
35. Coates (2017) makes this argument.

References

- Aristotle. 2002. *Nicomachean Ethics*. Translated by Christopher Rowe with commentary by Sarah Broadie, New York, NY: Oxford University Press.
- Arpaly, Nomy. 2000 "Rationally Acting against One's Better Judgment." *Ethics* 110: 488–513.
- Baier, Annette. 2010. *Reflections on How to Live*. New York, NY: Oxford University Press.
- Bakewell, Sarah. 2011. *How to Live: Or a Life of Montaigne in One Question and Twenty-one Attempts at an Answer*. New York, NY: Other Press.
- Barrett, Lisa Feldman. 2017. *How Emotions Are Made: The Secret Life of the Brain*. New York, NY: Houghton Mifflin Harcourt.
- Benbaji, Hagit. 2013. "How is Recalcitrant Emotion Possible?" *Australasian Journal of Philosophy* 91: 577–599.
- Brady, Michael. 2009. "The Irrationality of Recalcitrant Emotions." *Philosophical Studies* 145: 413–430.
- Cassam, Quassim. 2015. *Self-Knowledge for Humans*. New York: Oxford University Press.
- Coates, D. Justin. 2017. "A Wholehearted Defense of Ambivalence." *Journal of Ethics* 21: 419–444.

- Confucius. 2003. *Analects*. Translated by Edward Slingerland, Indianapolis, IN: Hackett Publishing.
- Crone, Katja. 2020. "The Self-Understanding of Persons beyond Narrativity." *Philosophical Explorations* 23: 65–77.
- D'Arms, Justin and Daniel Jacobson. 2000. "The Moralistic Fallacy: On the 'Appropriateness' of Emotions." *Philosophy and Phenomenological Research* 61: 65–90.
- . 2003. "The Significance of Recalcitrant Emotion (or Anti-Quasijudgmentalism)." *Royal Institute of Philosophy Supplement* 52: 127–145.
- Döring, Sabine. 2015. "What's Wrong with Recalcitrant Emotions? From Irrationality to Challenge of Agential Identity." *Dialectica* 69: 381–402.
- Fischer, John Martin and Neal A. Tognazzi. 2009. "The Truth about Tracing." *Nous* 43(3): 531–556.
- Goleman, Daniel. 1995. *Emotional Intelligence: Why it Can Matter More than IQ*. London: Bloomsbury.
- Greenspan, Patricia. 1988. *Emotions and Reasons: An Inquiry into Emotional Justification*. New York: Routledge, 1988.
- Harper, Aaron. 2015. "Nietzsche's Thumbscrew: Honesty as a Virtue and Value Standard." *Journal of Nietzsche Studies* 46(3): 367–390.
- Helm, Bennett. 2001. *Emotional Reason: Deliberation, Motivation, and the Nature of Value*. Cambridge: Cambridge University Press.
- . 2015. "Emotions and Recalcitrance: Reevaluating the Perceptual Model." *Dialectica* 69: 417–433.
- Hume, David. 2007. *A Treatise on Human Nature*. Edited by David Fate Norton and Mary J. Norton. New York: Oxford University Press.
- Hutto, Dan. 2016. "Narrative Self-Shaping: A Modest Proposal." *Phenomenology and the Cognitive Sciences* 15: 21–41.
- Lamb, Roger. 1987. "Objectless Emotions." *Philosophy and Phenomenological Research* 48: 107–117.
- Mele, Alfred. 1989. "Akritic Feelings." *Philosophy and Phenomenological Research* 50: 277–288.
- Moran, Richard. 1988. "Making up Your Mind: Self-Interpretation and Self-Constitution." *Ratio* 1(2): 135–151.
- Price, Carolyn. 2006. "Affect without Object: Moods and Objectless Emotions." *European Journal of Analytic Philosophy* 2: 49–68.
- Pugmire, David. 1998. *Rediscovering Emotion*. Edinburgh: Edinburgh University Press.
- Räikkä, Juha. 2005. "Irrational Guilt." *Ethical Theory and Moral Practice* 7: 473–485.
- Rorty, Amelie. 2010. "A Plea for Ambivalence." In *The Oxford Handbook of Philosophy of Emotion*, edited by Peter Goldie, 425–445. New York, NY: Oxford University Press.
- Rudd, Anthony. 2009. "In Defense of Narrative." *European Journal of Philosophy* 17(1): 60–75.
- Scaife, Robin. 2014. "A Problem of Self-Knowledge: Taking the Implications of Confabulation Seriously." *Acta Analytica* 29: 469–485.
- Shklar, Judith. 1984. *Ordinary Vices*. Cambridge, MA: Harvard University Press.

- Shoemaker, David. 2011. "Attributability, Answerability, and Accountability: Toward a Wider Theory of Moral Responsibility." *Ethics* 121: 602–632.
- Smith, Angela. 2004. "Conflicting Attitudes, Moral Agency, and Conceptions of the Self." *Philosophical Topics* 32: 331–352.
- . 2005. "Responsibility for Attitudes: Activity and Passivity in Mental Life." *Ethics* 115: 236–271.
- . 2008. "Control, Responsibility, and Moral Assessment." *Philosophical Studies* 138: 367–392.
- Solomon, Robert C. 1993. *The Passions: Emotions and the Meaning of Life*. Indianapolis, IN: Hackett Publishing.
- . 2003. *Not Passion's Slave: Emotions and Choice*. New York, NY: Oxford University Press.
- . 2007. *True to Our Feelings: What Our Emotions Are Really Telling Us*. Oxford: Oxford University Press.
- Stocker, Michael. 1996. *Valuing Emotions*. Cambridge: Cambridge University Press.
- Strawson, Galen. 2004. "Against Narrativity." *Ratio* 17: 428–542.
- Szanto, Thomas. 2017. "Emotional Self-alienation." *Midwest Studies in Philosophy* 41: 260–286.
- Taylor, Charles. 1985. "Self-interpreting Animals." In *Human Agency and Language: Philosophical Papers*, Vol. 1, Cambridge, MA: Harvard University Press.
- . 1989. *Sources of the Self*. Cambridge, MA: Harvard University Press.
- Thomason, Krista K. 2018. *Naked: The Dark Side of Shame and Moral Life*. New York, NY: Oxford University Press.
- . 2022. "How Should We Feel About Recalcitrant Emotions?" In *Self-Blame and Moral Responsibility*, edited by Andreas Brekke Carlsson, 134–148. New York, NY: Cambridge University Press.
- Ware, Owen. 2009. "The Duty of Self-Knowledge." *Philosophy and Phenomenological Research* 79: 671–698.
- Windsor, Mark. 2019. "What is the Uncanny?" *British Journal of Aesthetics* 59: 51–65.
- Wolf, Susan. 2015. "Character and Responsibility." *The Journal of Philosophy* 112: 356–372.
- Zahavi, Dan. 2007. "Self and Other: The Limits of Narrative Self-Understanding." *Royal Institute of Philosophy Supplement* 60: 179–202.