UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

EMOTIONS, REASONS, AND RATIONALITY

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

KELLY EPLEY
Norman, Oklahoma
2016

EMOTIONS, REASONS, AND RATIONALITY

A DISSERTATION APPROVED FOR THE
DEPARTMENT OF PHILOSOPHY

BY

_____
Dr. Wayne Riggs, Chair

_____
Dr. Sherri Irvin

_____
Dr. Amy Olberding

_____
Dr. Linda Zagzebski

_____
Dr. Jill Irvine

Dedication

For my family

Acknowledgments

I owe a debt of gratitude to many for help and guidance in my career as a philosopher. Thanks especially to my parents, grandparents, and siblings. They taught me how to be a student, supported me through many years of school, and never doubted my abilities even though I sometimes did. Thanks to Patrick Epley for going through the process of getting a Ph.D. with me, for keeping me on task, and for all of the help and encouragement.

Thanks to my committee members, Sherri Irvin, Amy Olberding, Linda Zagzebski, and Jill Irvine. Each has contributed something important to my philosophical development. Thanks especially to my advisor Wayne Riggs for guiding me through the whole process, and for always being upbeat, even with the occasional setbacks.

Thanks to Steve Ellis, Shelley Konieczny, and Gabe Serrano for making sure all of my paperwork got where it needed to go, and for generally keeping me on track to finish my degree.

Thanks also to the whole OU philosophy department. It's been a fantastic place to learn philosophy.

Table of Contents

Abstract

Intuitively, emotion/judgment conflicts are irrational because they are disruptive. They create disturbances and divisions in our generally well-ordered and cohesive selves. When you fear the spider, even though you judge that it isn't harmful to humans, you may recoil from its frightful appearance. You might also spend your mental and practical resources and devising a safe disposal of the pest with a long broom handle that keeps a comfortable distance between you and it, even though your judgment implies that such precautions aren't necessary. Unless or until the conflict is resolved, each element of the conflict vies for influence over your thoughts and actions. In this dissertation, I give an account of emotion/judgment conflicts according to which they are conflicts of assent-bearing attitudes. As attitudes, emotions and judgments both entail rational commitments, at a minimum, to refrain from having inconsistent attitudes. Emotion/judgment conflicts violate these commitments.

A major objection to this view is that emotions appear to be inconsistently responsive to reasons that support altering or extinguishing it. Some emotions easily change in response to reflection, but some are highly recalcitrant. Therefore, emotions don't obviously have the sort of relationship to reasons and reflection that our attitudes have. Attitudes are reason responding. In response to this objection, I argue that emotions do generally respond to reflection on our reasons, even in some cases in which they appear to be recalcitrant because they are unmoved by the reasons that formed the basis of our judgment. The modes of reflection that would be able to adjudicate the conflict are just the sorts that our emotional faculties respond to: attention redirection, information elaboration, and imagination. These modes of reflection can correct for

hasty and myopic judgments as well as ill-founded emotions by widening our reasons set and recalibrating our weighting schemes.

An important advantage of the attitude view of emotion/judgment conflicts over alternatives, I argue, is that the alternatives can't do justice to our emotional faculties. Other views generally explain the irrationality of emotion/judgment conflicts by appealing to the rational primacy of our judgment faculty. Emotions that persist in the face of a conflicting judgment are assumed to be without reasons to support them and judgments are presumed the rightful guides of thought and action. I argue that our emotion and judgment faculties can each be well-regulated, carefully monitored, and broadly sensitive to reasons in some domains, and each can produce outputs that are ill-founded, short-sighted, and mistaken in other domains. Our judgment faculty should not be given presumptive weight in cases in which our emotion faculties are of the former sort and our judgments of the latter. The attitude view can accommodate the possibility that the judgment requires revision rather than the emotion in some emotion/judgment conflicts.

.

Chapter 1

The Head and the Heart

In emotion/judgment conflicts, we experience discord, metaphorically speaking, between the head and the heart. These conflicts come in many familiar forms. You feel guilty for some harm you caused to another though you judge that you couldn't have done otherwise. You judge that there is no significant danger, but you're afraid nonetheless. You judge that you ought to be proud of your accomplishments, but you feel ashamed instead. You judge that you ought to be happy with your career, but you experience a profound and persistent sense of unhappiness and dissatisfaction.

Sometimes a conflict arises and resolves itself quickly. You recognize that your emotional impression of a situation formed before all the information was in, and the emotion dissolves or adjusts to accord with the new inputs. Imagine it's your first time experiencing turbulence on a plane and you're afraid. You assure yourself that it's nothing to worry about—that turbulence is the most likely explanation of the plane's jarring motions—even as you feel yourself trembling in your seat. Eventually, your fear dissipates as you reach a sense of certainty that the plane is holding a steady altitude and you aren't, for the moment, in any danger of crashing.

Other emotion/judgment conflicts persist. A conflict arises and the state of cognitive dissonance it creates refuses to dissipate. Returning to your first time experiencing turbulence on an airplane, imagine that although you rehearse the arguments in favor of your considered judgment time and again, your emotional state is unmoved. Instead of reaching a sense of certainty that your plane is in no danger, you continue to tremble and worry. You feel each bump on your journey and every time

your fear is renewed. You try to distract yourself with a book, but your mind wanders again and again back to the thought that *planes crash*. Even though you remind yourself that the likelihood of a crash is very low, you can't seem to quell your fear.

Popular wisdom about the nature of these conflicts is divided. On the one hand, we speak of emotions as passions, describing them as waves of feeling that pass over us and sometimes overcome us. Emotions are not under the direct control of the judging mind, so they can move us to feel, think, and do things that we judge incorrect or unworthy. On the other hand, there is a familiar adage, "the heart has reasons that reason knows not," that suggests a more optimistic view of our emotions. It suggests that our emotions have intelligence—that they are sensitive to reasons that our reasoning faculty may fail to access.

There is some truth in each manner of speaking about our emotions. Our powers of volition over our emotions are limited. Our emotions can arise in response to situations faster than our slower judgment faculties are able, preparing us to deal with potentially significant and time-sensitive issues as swiftly and efficiently as possible. Consequently, they're not always attuned to features of a situation that could make an emotional response appropriate in one circumstance, but inappropriate in another similar, yet importantly different one. Our emotions are not, however, insensible to reasons. On the contrary, our emotions respond to the presence or absence of value and goal-relevant objects and states of affairs in our environments. They are, in large part, conditioned responses that reflect prior learning and experience about the sorts of things that are beneficial or harmful to us.

Our emotions play important roles in shaping both our immediate responses to our situations and our long term thinking, decision-making, and planning. Discord between our emotions and judgments can be a greatly disruptive form of mental conflict. For this reason, emotion/judgment conflicts are generally thought to be irrational.

The purpose of this dissertation is to account for *how* emotion/judgment conflicts are irrational. There are three broad possibilities. The first possibility is that emotions/judgment conflicts are attitude conflicts. To have two contradictory attitudes simultaneously is irrational. Proponents of this view maintain that that emotions are, or involve, intentional attitudes such as judgments or desires. The second possibility is that emotions either violate or cause one to violate norms of rationality without implicating emotions in attitude conflicts. Views of this sort are premised on the notion that emotions are not, and could not be, attitudes. Instead, emotions are categorized along with non-commitment oriented mental states such as thoughts, construals, or inclinations. The third possibility is that emotion/judgment conflicts are not, strictly speaking, irrational. Even if emotions might sometimes cause irrationality, they are not directly implicated in the irrationality they cause. This position tends to be associated with the view that emotions are non-cognitive.

In this dissertation, I defend a version of the first view: the attitude conflict view. The attitude conflict view is typically associated with views on which emotions are argued to be, or to necessarily involve, judgments. Though my view bears similarities with these views, I want to distance my own view from them for reasons I will discuss below.

Emotions are thought to involve judgments because it's a challenge to make sense of our emotions without reference to judgments or judgment-like attitudes. After all, if one truly had no emotion-relevant attitudes about an object or state of affairs, the elicitation of an emotional reaction would be mysterious. Anthony Kenny's fear of gambling case is a helpful illustration:

> If a man says that he is afraid of winning £10,000 in the pools, we want to ask him more: does he believe that money corrupts, or does he expect to lose his friends, or to be annoyed by begging letters, or what? If we can elicit from him only descriptions of the good aspects of the situation, then we cannot understand why he reports his emotion as fear and not as hope.[1]

We expect that someone with such a fear has some attitudes that make sense of it, or else we'd suspect that he's misunderstood the emotion category that he's applied. Fear is only sensible when there is apparently danger present. Without a danger-appraisal of some sort, we can hardly fathom what his fear consists in.

The trouble with the notion that emotions *necessarily involve judgments* is that we can offer possible explanations of the fear of gambling case without references to present judgments. Consider the following:

> Perhaps the man who fears winning £10,000 in the pools was raised in a strict teetotalling, antigambling household and was habituated to fear gambling and other "sinful" addictions for most of his life. He no longer judges that gambling corrupts in the way that he used to, and hence he denies having the sorts of beliefs or judgments that would explain his fear. Yet, he is afraid of what might happen to him if he wins a sizable profit.

His fear of winning at gambling doesn't involve his judgment faculty's present outputs. His fear is, instead, a vestige of his past experiences, beliefs, and associations. This does not show that he has no anti-gambling attitudes presently. On the contrary, I defend the

[1] Anthony Kenny, *Action, Emotion and Will* (New York: Routledge, 1963) 192.

view that his fear is an attitude. But to have this attitude does not require that he presently judges accordingly.

A common objection to the view that emotions *are judgments* is that the outputs of our emotion faculties and our judgment faculties don't always coincide. Emotions and judgments result from different mental processes that don't necessarily produce convergent outputs. Consequently, conflicts between the deliverances of our emotion faculties and our judgment faculties are not uncommon. Defenders of the view that emotions are judgments must concede that the judgments that issue from our emotion faculties are of a different kind than those that issue from our judgment faculties in order to account for such conflicts. Nevertheless, defenders of the view that emotions are judgments argue that emotions involve something evaluative—an appraisal—that makes an emotion very like a judgment, though it may be of a different kind. In the fear of gambling case, this suggests that even though the man denies having a judgment that makes sense of his emotion, his emotion implies that he appraises gambling to be dangerous, at least implicitly.

The most well-known defender of the view that emotions are judgments, Robert Solomon, argues that the comparison between emotions and judgments has heuristic value. Out of all of our existing mental concepts, emotions and judgments are most alike. He explains,

> Emotions as judgments are not necessarily (or usually) conscious or deliberative or even articulate but we certainly can articulate, attend to, and deliberate regarding our emotions and emotion-judgments, and we do so whenever we think our way into an emotion, 'work ourselves up' to anger, or jealousy, or love. The judgments of love, for instance, are very much geared to the perceptions we have of our beloved, but they are also tied to all sorts of random thoughts, day-dreaming, hints and associations with the beloved, with all sorts of memories and intentions and imaginings. A judgment may be made at a certain

> time, in a certain place ('I loved you the first time I ever saw you') but one continues to make, sustain, reinforce, and augment such judgments over an open-ended amount of time.[2]

Here, Solomon enumerates some of the features that emotions and judgments share. In particular, he notes that emotions and judgments both come in non-reflective and reflective varieties and momentary and sustained varieties. Most importantly, they are the sort of thing that we *can* "articulate, attend to, and deliberate" about, though we often don't.

As attitudes, emotions and judgments do share some similar features. The heuristic value of the comparison is, however, counterbalanced by the number of objections it invites. Since emotions are not judgments, there are plenty of disanalogies to point to. Most commonly, it's pointed out that (1) evaluative judgments can be dispassionate while emotions cannot,[3] and (2) judgments are under our executive control, meaning we're able to consciously and deliberately manipulate their contents, while emotions are not.[4]

Departing from judgment views, I argue that emotions are unique intentional attitudes. While they do hold similarities with judgments insofar as they are both

---

[2] Robert C. Solomon, "Emotions, Thoughts, and Feelings: Emotions as Engagements with the World," in *Thinking about Feeling: Contemporary Philosophers on Emotion*, edited by Robert C. Solomon (Oxford: Oxford University Press, 2004) 83.
[3] Jerome A. Shaffer, "An Assessment of Emotion," *American Philosophical Quarterly* 20.2 (1983): 206; Jenefer Robinson, "Emotions: Biological Fact or Social Construction," in *Thinking about Feeling: Contemporary Philosophers on Emotion*, edited by Robert C. Solomon (Oxford: Oxford University Press, 2004) 29-30; Peter Goldie, *The Emotions: A Philosophical Exploration* (Oxford: Oxford University Press, 2002) 18-19.
[4] Jesse J. Prinz, *Gut Reactions: A Perceptual Theory of Emotions* (Oxford: Oxford University Press, 2004) 38-9; Robert C. Roberts, *Emotions: An Essay in Aid of Moral Psychology* (Cambridge: Cambridge University Press, 2003) 84.

intentional attitudes that involve evaluation and assent, the comparison to judgment goes no farther. Consequently, my view has some very important advantages over the judgment view. In particular, I'm not committed to defending an analogy between mental states that have different causes and separate functions. Instead, I can focus on the uniqueness of emotions, their architecture, and the roles they play in our lives. Additionally, I'm not committed to defending anything more than a modest cognitivist thesis. Cognitivism is, minimally, the view that emotions are or involve cognitions (i.e. evaluations or appraisals) at least most of the time. Judgment views are typically tethered to stronger cognitivist commitments according to which each emotional state is necessarily preceded by a token of a particular type of cognition. It's these initiating cognitions that enable judgment theorists to compare emotions to judgments. Non-cognitivists point to research suggesting that some emotions are initiated without preceding cognitions. If this is true, it is a major blow to the judgment view. It does little damage, by contrast, to the attitude view.

On the attitude view that I defend, emotions are complex processes that manifest in extended, recursive emotional episodes. To undergo these processes is to be in an emotional state. To have a stable tendency to be in certain states with respect to a certain stimulus or class of stimuli is an emotional disposition. Cognitions play a crucial role in the formation of emotional dispositions. Cognitions are also involved in our emotional states. In the vast majority of cases involving persons who have progressed beyond human infancy, cognition plays a role in the initiation of an emotional state. In those cases in which it does not, cognitions will play a role in the ensuing emotional episode with few, if any, exceptions after the earliest stages of human development.

To say that an emotion is an attitude is to say that we have reasons for them and that they entail rational commitments that we violate when we experience emotion/judgment conflicts. The view that emotions are attitudes, I argue, provides the most plausible account of the irrationality of emotion/judgment conflicts. It also offers the most reasonable account of how responsible agents ought to respond to emotion/judgment conflicts. The alternative views suggest that the emotion is always the source of the irrationality implicated in emotion/judgment conflicts. I argue to the contrary that the *attitude conflict* is the source of irrationality; when conflicts arise, each element of the conflict is a candidate for correction in resolving the conflict. Our judgment faculty is susceptible to errors resulting from defective weighting schemes or too narrow focus. Sometimes, on reflection, we discover that our emotions are better supported by our reasons than our judgments. Emotions are thus a useful corrective for poor judgment and we should not assume that judgment always takes precedence over emotion in a conflict.

One reason why emotions are assumed not to be reason responding is that they don't always respond to our current judgments about what's appropriate or correct. I argue that responding to *judgment* is not a necessary requirement for being a reason responding attitude. Judgment oriented reasoning is one element of a larger collection of reflective tools and capabilities that we can use to reflect upon our reasons and shape our attitudes. Our emotions don't necessarily respond to our current judgments, particularly our deliberative judgments, because our deliberative judgments are formed in response to information that is presently accessible and seems salient to us. Our emotions, by contrast, embody a history of past experiences, attitudes, and emotions in

implicit memory that shape the way we experience emotions in the present. Still, we are capable of shaping our emotions in the present with reflection. Emotions respond to reflective activities such as attention redirection, information elaboration, and imagination. These activities can, when used effectively, interrupt the processes that keep an inappropriate emotion on its course. This provides us with opportunities to draw out more appropriate ones according to our reasons.

It is, of course possible for an emotion to be genuinely recalcitrant, and therefore beyond our influence or control. This is hardly a reason to deny attitude status to emotions and cede emotional helplessness. Every variety of attitude admits of such cases, including judgments. Also, as with other varieties of attitude, genuine recalcitrance tends to be the exception rather than the norm for our emotions. Ordinarily, we are capable of shaping and redirecting our emotional impulses according to our reasons, even though it may require considerable effort in some cases. I take up these issues regarding reasons and reflection in chapter 3.

Of course, the view that emotions are attitudes implies that emotion/judgment conflicts are attitude conflicts. This implication invites objections. There are two major objections to the attitude conflict view. First, it's objected that the kind of irrationality implied by an attitude conflict is inordinately serious. If we grant that emotions and judgments are both attitudes with evaluative contents, then we're committed to the view that emotion/judgment conflicts involve attitudes with contradictory evaluative contents. This has the appearance of a very serious form of irrationality. It's very close to judging A and not A at the same time, which is a violation of the law of noncontradiction. Peter Goldie has joined a number of philosophers in protesting, "we

just do not believe contradictions."[5] Nor, arguably do we make simultaneous judgments that are contradictory. Accordingly, it's argued that even if emotion/judgment conflicts are irrational, they aren't as serious as this.

The second objection is of a piece with the first. It's objected that to attribute irrationality to a person who is experiencing an emotion/judgment conflict violates the principle of charity. The principle of charity enjoins us to avoid attributing irrationality to a person if it's possible to. We can avoid it by interpreting emotions as non-attitudes.

I answer these objection in chapters 4 through 6. In response to the objection that attitude conflict view implies that emotion/judgment conflicts are more irrational than they really are, I argue that emotion/judgment conflicts fall on a spectrum of the seriousness of the irrationality they represent. The most familiar kinds of conflict do not fall on the most serious end of the spectrum. It's worth noting that factors that make a violation serious sometimes also make us less culpable for them. We are most culpable for conflicts that we are capable of resolving with reflection. However, when we have emotions or other attitudes that are not capable of being moved by reflection, we bear less blame for them.

In response to the objection that attributions of irrationality in cases of emotion/judgment conflict are uncharitable, I argue that there are sufficient reasons to attribute irrationality to a person who is exhibiting an emotion/judgment conflict given how disruptive emotion/judgment conflicts can be. The principle of charity instructs us

[5] Peter Goldie, "Getting Feelings Into Emotional Experience in the Right Way," *Emotion Review* 1.3 (2009): 237.

to avoid attributions of irrationality when it can be avoided, but emotion/judgment conflicts are disruptive in a way that makes interpretations of irrationality appropriate.

In sum, the view I offer explains how emotions function as reason-responding attitudes that can conflict with our other reason-responding attitudes, and are thus implicated in irrational attitude-conflicts when they conflict with our judgments. Alternative accounts of the irrationality of emotion/judgment conflicts assume that there are norms of rationality that privilege judgment as the rational choice in an emotion/judgment conflict. These accounts don't appropriately acknowledge the intelligence of emotions or the opportunities they provide us to correct errant judgments resulting from inappropriately narrow focus or faulty weighting schemes. This, I contend, downplays the importance of emotions in our mental lives. Emotion/judgment conflicts are disruptive, but in most cases we shouldn't jump to the conclusion that the emotion is in error. The conflict afford us the opportunity and occasion to reflect on our reasons and revise our attitudes according to what we have most reason to hold.

Chapter 2

Literature Review

In this chapter, I introduce the concept of an emotion, survey contemporary theories of emotion in contemporary psychology and philosophy, and comment on the ramifications of these views for how we ought to understand the nature of the irrationality that emotion/judgment conflicts are apparently implicated in. At the end of the chapter, I present a view of emotions and a preliminary sketch of the account of emotion/judgment conflicts that I defend in the dissertation.

1. What are emotions?

Emotions are responses to value-relevant objects and states of affairs: Our health and well-being, the health and well-being of loved ones, our dignity, our valued possessions, our plans, our good social standing, and our moral self-conception. Successes, gains, advantages, threats, losses, obstacles, antagonists, and so forth elicit our emotions.

As we encounter value relevant objects in the world or in our imaginations, our emotional experiences—the raw phenomena that we try to account for with an analysis of the concept of an emotion—are episodic, with multiple components. The various components of an emotional episode emerge in dynamic stages that unfold, recur, and adjust as new information is taken in. In a typical emotional episode there is an appraisal component, a bodily change component, an action tendency component, a cognitive style component, and a feeling component. For illustrative purposes, imagine

that you are alone in your house late at night and you are startled from your repose by a strange noise, followed by a steady thumping that resembles footsteps.

i. The *appraisal component* of an emotional episode is comprised of a cognition or perception which determines the situation to be emotionally relevant and how.

   When you hear what sounds to you like footsteps in your house at night and experience fear, you recognize (perhaps implicitly) that your situation—or at least ones like it—represents a potential threat to your wellbeing. Insofar as wellbeing is a long-term goal or value for you, you'll be disposed to appraise situations that put it in jeopardy negatively.

ii. The *bodily change component* consists in changes that occur in the autonomic nervous system (affecting heart, lungs, stomach, and other viscera), the motor system (affecting posture and facial expression), and the endocrine system (regulating hormones and chemicals that affect the body and brain).[6]

   As you lie in bed listening to the footsteps, your eyes widen, your muscles freeze, and your heart begins to pound, among other changes. These changes may or may not rise to the level of consciousness.

[6] Antonio Damasio, *Descartes' Error: Emotion, Reason, and the Human Brain* (New York: Putnam, 1994) 183; Joseph E. LeDoux "Rethinking the Emotional Brain" *Neuron* 73 (2012): 662-663.

iii. The *action tendency component* of an emotion consists of the preparations one undergoes to respond to the situation advantageously such as approaching, avoiding, rejecting, attending, and resting.[7] These responses can be either innate or learned.

The characteristic action-tendencies for fear include avoidance and inhibition, which promote self- and value-preservation.[8] Thus, as you experience intruder-fear listening to the sound of footsteps, you'll likely lie very quiet and still in your bed at first. This helps you to gain information about the nature of the potential threat. It also helps you to avoid drawing attention to your presence, which prevents a premature confrontation. Then, if the threat does not dissipate, you may begin to make preparations for a possible encounter with a prowler by hiding, assuming a posture that you learned in a self-defense class, or reaching for a weapon.

iv. The *cognitive style component* is comprised of emotion-relevant attention biases and emotion-congruent thinking.

[7] These are just a few examples from the list in Nico H. Frijda, "Emotion, Cognitive Structure, and Action Tendency," *Cognition and Emotion* 1.2 (1987): 133.
[8] Ibid., 135.

Emotion-relevant objects may grab and hold our attention without a sustained effort in a way that other objects do not.[9] Emotions also affect the global or local quality of our attentional focus. Positive emotions tend to produce global biases in global-local visual processing tasks. A global bias facilitates attention to large-scale, or "big picture" qualities, of one's environment. Negative emotions produce local biases that facilitate attention to individual components of one's environment.[10] In fear, you focus on particular elements of threat and coping potential in your environment. In the intruder-fear scenario, you'll likely be gripped by the suspicious sounds that initiated your fear. It may be more difficult than usual to willfully turn your attention to other non-threat-related matters, such as the book you were reading when you first heard footsteps. You might instead feel compelled to attend to escape routes and heavy objects that could serve as make-shift weapons, as well as unfamiliar sounds.

Emotions are also typically accompanied by characteristic patterns of thought. Fear encourages recollections, imaginings, and judgments related to escape and defense. You might recall your premeditated emergency plans and form a judgment about their fittingness to your current situation. You might imagine possible outcomes of your scenario on the basis of the ideas you have about what happens in home invasions from the media or other sources.

[9] Michael S. Brady "Emotion, Attention, and the Nature of Value," in *Emotion and Value*, edited by Sabine Roeser and Cain Todd (Oxford: Oxford University Press, 2014) 53-56.
[10] Ibid., 57-58.

v. The *feeling component* of an emotion is comprised of the subject's felt experience of one or more of the other components.

   You may feel yourself trembling and undergoing other bodily changes as you listen for the intruder. Your feelings may also be composed of your experience of the attention biases and the elements of memory, thought, and imagination that arise in a fear episode.

As your emotion unfolds, you take in new information, which can produce changes in the course of the emotion's progression. If you begin to suspect that the footsteps you thought you were hearing are actually the sounds of acorns falling on a wooden deck on the other side of your house, this will cause changes to occur in the components of your emotional state. If you become convinced that there is no intruder, it may inhibit your fear-state altogether, permitting you to return to the book you were reading with a sigh of relief. However, if you aren't convinced that the noise you hear is acorns because you hear what sounds like whispering voices along with the gentle thumping sounds of apparent footsteps, your fear will heighten instead. The potential for danger has become even more apparent and this is reflected in the character that your emotional experience takes on as it progresses. Instead of returning to your book, your fear intensifies as you prepare yourself for the worst.

2. Contemporary psychological theories of emotion

For the purposes of this dissertation, I organize contemporary psychological theories of emotion into five categories: body/feeling theories, two-component theories, affect program theories, appraisal theories, and neural network theories.[11] These theories are distinguished by the manner in which they parse the emotional episode—how they conceive of the order and priority of its components. Does the appraisal component come first? Or is it typically preceded by the bodily change component? How are the other components related to these in a chain of cause and effect?

i. Body/feeling theory

William James's (1884) influential theory of emotion prioritizes the bodily change and feeling components of an emotion episode, identifying emotions with the felt experience of bodily changes. James suggests that an emotion is an experience of the body's response to a stimulus. In an emotional episode, the sensory cortex is activated by a stimulus, which results in somatic and motor responses. These bodily responses result in feedback to the sensory cortex, producing emotional feelings.[12] The other components of an emotional episode are largely overlooked on this theory. To explain the differentiation of particular emotions such as anger, joy, and sadness, James

[11] This grouping of categories is roughly based on the list in Agnes Moors, "Theories of Emotion Causation: A Review," in *Cognition and Emotion: Reviews of Current Research and Theories*, edited by Jan De Houwer and Dirk Hermans (New York: Psychology Press, 2010) 1-37.
[12] Here, I follow Moors' characterization of the theory. Ibid., 10.

hypothesizes that each emotion-type has a unique profile of bodily response patterns that gives it its distinctive quality.[13]

More recently, Antonio Damasio has defended a view that resembles James' theory in some important respects. Damasio (1994) argues that emotions are coordinated changes in the body caused by a mental evaluative process (an appraisal) and followed by associated feelings and cognitive styles.[14] Emotional feelings, on Damasio's view, are brain states that track coordinated changes in the body while juxtaposing these changes with the mental images which initiated them.[15] His research also indicates that some emotional feelings do not involve *actual* bodily changes. Instead, they engage an "as-if" loop which simulates bodily changes in the brain without enacting them. This results in a "fainter image of an "emotional" body state, without having to reenact it in the body proper."[16] Thus Damasio's theory, like James's, identifies coordinated change in the body (or simulated changes) as an essential component of an emotion. Unlike James's theory, Damasio conceives of felt experiences of those changes as a non-essential emotional effects. Sometimes the coordinated changes in the body that constitute an emotion are registered in the brain unconsciously.

---

[13] William James, "What is an Emotion?" in *What is an Emotion? Classic Readings in Philosophical Psychology*, edited by Cheshire Calhoun and Robert C. Solomon (Oxford: Oxford University Press, 1984) 131.
[14] Damasio, *Descartes Error,* 139, 147.
[15] Ibid., 144-8.
[16] Ibid., 155.

ii. Two-component theory

James's theory assumed that emotions could be differentiated by unique patterns of bodily responses that produce emotional experiences. These patterns are identifiable with particular emotions such as anger and happiness. Schacter and Singer (1962) expressed doubt that differentiated patterns of bodily response determine which emotion a person experiences.[17] They hypothesized that without cognitive labeling, the feeling component of an emotional experience would be undifferentiated affect. Mandler helpfully describes the two-component theory in terms of a jukebox analogy: affect is the moving turntable and the cognitive label is the vinyl record. Both are necessary to make an emotion—the tune.[18]

To support their hypothesis, Schacter and Singer designed an experiment to show that they could induce an undifferentiated state of arousal in subjects and influence which emotion that the subject reported experiencing by manipulating the situational context. In the experiment, they induced a state of arousal in subjects using an injection of epinephrine. Epinephrine mimics a discharge of the sympathetic nervous system, producing symptoms such as tremor, palpitations, flushing, and accelerated

[17] There is a helpful discussion of two-component theories in Moors, "Theories of Emotion Causation," 11-12. Moors notes that Barrett's conceptual act theory is sometimes categorized as a two-component theory. On Barrett's view, emotions are composed of a "core affect," or bodily response component, and a categorization component. On Barrett's theory, these components do not occur in succession. Instead, they are simultaneous, and thus categorization influences the emotional experience. Categorization is achieved by a perceptual process rather than a conscious cognition. Importantly, on Barrett's view, individual emotions are not natural kinds, they are socio-cultural constructions. The perceptual component of an emotional experience is shaped by one's conceptual categories and one's prior knowledge and experience. Ibid., 21-24.

[18] See George Mandler, *Mind and Emotion* (New York: Wiley, 1975). Mandler is a two-component emotion theorist.

breathing. Participants were told that the injection was a vitamin compound and that the experimenters were testing the effects of the compound on vision. A control group received a placebo in the form of a saline injection, which produces no side effects.[19]

One group of participants were given an accurate description of the side effects they would experience as a result of the injection. A second group of participants were given a misleading description of the side effects that they might experience: headache, itching, numb feet. A third group received no information about the injection's side effects.[20] Subsequent to the injection, participants were placed in a room with a confederate of the experiment who engaged in either euphoric behaviors in one condition or angry behaviors in another. Experimenters measured participants' emotions by observing their behaviors from behind a one-way mirror and recording self-reports. Participants who were not provided with an appropriate explanation of their symptoms—either because they were given no information about the injection's side effects or because they were given misleading information—exhibited an emotion that matched the circumstance that experimenters placed them in. If they were placed with the euphoric confederate, they tended to exhibit happy emotions, and if they were placed with the angry confederate, they tended to exhibit angry emotions. This suggests that the emotions of these participants were shaped by the cognitive labels, happiness and anger, that were made readily available to them by the presence of the confederate. Participants who were given correct information about their symptoms neither displayed

[19] Stanley Schacter and Jerome A. Singer, "Cognitive, Social, and Physiological Determinants of an Emotional State," in *What is an Emotion? Classic Readings in Philosophical Psychology*, edited by Cheshire Calhoun and Robert C. Solomon (Oxford: Oxford University Press, 1984) 178.
[20] Ibid.

nor reported happy or angry emotions, presumably because they had a non-emotional explanation of their symptoms ready to hand.[21]

On the basis of their findings, Schacter and Singer developed a theory on which emotions have two essential components: (1) bodily arousal produced by a stimulus, and (2) a subsequent conscious cognitive labeling of the causes of one's arousal. The act of conscious labeling, they suggest, has a "steering function." They explain, "Cognitions arising from the immediate situation as interpreted by past experience provide the framework within which one understands and labels his feelings. It is the cognition which determines whether the state of physiological arousal will be labeled as "anger," "joy," "fear," or whatever."[22]

iii. Affect program theory

Affect program theory studies the neural signatures of emotions. According to affect program theorists, there are a limited number of "basic" emotions with unique neural signatures that produce specific sets of bodily responses and experiences. Ekman, for example, postulates that enjoyment, anger, sadness, fear, disgust, and surprise are basic emotions with unique neural circuits based on his research on patterns of emotional facial expressions.[23] Views about the nature of the triggers that produce emotions are typically borrowed from other compatible theories which offer accounts of

[21] Ibid., 181-2.
[22] Ibid., 174.
[23] Paul Ekman, "An Argument for Basic Emotion," *Cognition and Emotion* 6.3/4 (1992): 172-4.

emotion elicitation such as appraisal theories and neural network theories (discussed below).[24]

Though there are limited number of basic emotions on these views, each basic emotion admits of numerous variations in intensity, situational antecedents, and other variables. These variations are supposed to account for much, if not all, of our emotional vocabulary. However, some affect theorists have taken note of some apparent exceptions such as shame, guilt, contempt,[25] and pride that are non-basic on most models, but do not obviously have other basic emotions as their core components.[26]

iv. Appraisal theory

Appraisal theorists hold that emotional episodes begin with an appraisal. An appraisal is an evaluation of the personal significance of a stimulus. It is generally thought to be cognitive, but it need not be conscious. Once an appraisal is elicited, action tendencies, changes in cognitive style, and changes in the body are initiated. These are combined in emotional experience to produce the feeling component.[27]

---

[24] See Moors, "Theories of Emotion Causation," 20-21.
[25] Carroll E. Izard, "Forms and Functions of Emotions: Matters of Emotion-Cognition Interactions," *Emotion Review* 3.4 (2011): 376.
[26] Jessica L. Tracy and Daniel Randles, "Four Models of Basic Emotions: A Review of Ekman and Cordano, Izard, Levenson, and Panksepp and Watt," *Emotion Review* 3.4 (2011): 400. Pride, for example, has some unique features that don't reduce to other basic emotions. This could be evidence that it is itself a basic emotion. If it is a basic emotion, however, it diverges from basic emotion models on existing theories. Pride involves high level neurological mechanisms that other basic emotions don't.
[27] Agnes Moors, "Flavors of Appraisal Theories of Emotion," *Emotion Review* 6.4 (2014): 305.

There are a number of appraisal variables that determine when an emotion is elicited, which one it is, and how intense it is. The two variables that most appraisal theorists agree upon are goal relevance and goal congruence.

a. Goal relevance: A stimulus indicates something about the status of one's goals. The more important the goal, the more important changes in its status will be for a person, and the more intense emotion elicited in response to significant changes to the status of this goal will potentially be.

b. Goal congruence: Stimuli that are goal congruent elicit positive emotions. Stimuli that are goal incongruent elicit negative emotions.

Other commonly postulated appraisal variables include certainty, coping potential, and agency/blame. [28]

Smith and Lazarus proposed the term "core relational themes" to describe the combination of appraisal subcomponents corresponding to the above appraisal variables that are involved in particular emotions. In sadness, for example there is an appraisal of "irrevocable loss." "Irrevocable loss" is a core relational theme involving several appraisal subcomponents: "(1) a goal commitment [goal relevance] as being (2) threatened or harmed [goal congruence] in a manner that (3) cannot be avoided or repaired [coping potential]."[29]

While Smith and Lazarus contend that the appraisals that cause emotions are multidimensional, Frijda suggests that they are, in some of their manifestations, fairly simple. The complexity may be added via cognitive elaboration after the emotion has been triggered due to the activation of an emotion-*schema* or emotion *script*. This might

[28] Moors, "Theories of Emotion Causation," 15-16.
[29] Craig A. Smith and Richard S. Lazarus, "Appraisal Components, Core Relational Themes, and the Emotions," *Cognition and Emotion* 7.3/4 (1993): 236. The brackets indicating the relevant appraisal variables are my additions.

better explain, for example, why we occasionally get angry at non-agents such as inanimate objects, but experience them, in our anger, as if they were agents capable of carrying blame. You stub your toe on a rock, and in anger, you violently kick the rock into a pit as if to punish it. The appraisal that initiated the emotion may have been simply that you've experienced "acute goal frustration." It's your anger script, which contains agency and blame components, that facilitates the attributions you make to the stone.[30]

An appraisal can be elicited by either a rule-based mechanism or an associative one. An emotion is elicited by a rule based mechanism if the value of the various appraisal variables are computed from propositional inputs and a corresponding emotion selected. An emotion is elicited by an associative mechanism if past appraisals stored in memory are retrieved or reinstated by perceptual inputs. Some appraisal theorists add a third mechanism: innate sensory-motor connections. A few stimuli such as faces and loud noises are believed to be innate emotional triggers that activate sensory motor connections. Inputs of innate sensory-motor connections are simple perceptions that have not been integrated into a perceptual representation. It is, however, controversial to call this third mechanism a mechanism of appraisal if the assumption is that appraisals are cognitive, as I discuss below.[31]

[30] Nico H. Frijda, "The Place of Appraisal in Emotion," *Cognition and Emotion* 7.3/4 (1993): 364, 373.
[31] See Moors, "Theories of Emotion Causation," 17-18.

v. Neural network theory

Neural network theories draw from work in the conditioning and the semantic network models of memory literatures. As infants, our earliest emotions are unconditioned, and they are elicited by a limited range of biologically relevant stimuli. As we develop, the range of stimuli capable of evoking emotions is expanded and elaborated through conditioning. Beyond infancy, emotions are caused by the activation of emotion schemata recorded in memory.[32]

Early neural network theory models favored a computational theory of mind. On these models, emotional memories are organized into emotion schemas. Individual information components—information regarding the stimulus that brought about the emotion, action tendencies, resulting behaviors, and conceptual meanings of the emotional situation—are encoded in distinct memory nodes that are associated with the emotion schema. Additional stimuli become associated with an emotion schema when a non-emotion-eliciting stimulus co-occurs with an emotional one. Also, we make generalizations from our present associations when we encounter a stimulus that is similar, though not identical to an emotional stimulus.[33] Schema activation can occur in a number of ways. One is stimulus recognition. Another is the operation of sensory-motor processes associated with the memories encoded in the schema. These modes of schema activation are called "entry points."[34]

---

[32] Ibid., "Theories of Emotion Causation," 18.
[33] Howard Leventhal and Klaus Scherer, "The Relationship of Emotion to Cognition: A Functional Approach to Semantic Controversy," *Cognition and Emotion* 1.1 (1987): 9.
[34] Moors, "Theories of Emotion Causation," 18-9.

Later neural network theories favor connectionist models of mind over the computational ones assumed by earlier theories. On these later theories, memory components are distributed. The do not occupy separate, distinct nodes. Also, network activation is not sequential. Multiple components can be activated simultaneously.[35]

3. Philosophical views of emotions

The philosophical views of emotions offer diverging accounts of how emotions fit into our landscape of cognitive and mental concepts. They are divided into cognitive attitude views, cognitive non-attitude views, and non-cognitivist views. The views outlined here provide different sets of parameters for answering the question of how emotion/judgment conflicts might be irrational.

i. Cognitivist attitude views

Cognitive attitude views posit that emotions are or involve attitudes, usually judgments.[36] Judgment is the attitude that most closely corresponds to the notion of an appraisal, which according to appraisal theories of emotion, initiates an emotional episode. To experience an emotion, on these views, is to make a judgment (or a set of judgments, or a complex judgment) about how things are going for you. Jealousy, for example, involves a judgment that a valued relationship is threatened by defection.

---

[35] Ibid., 20.

[36] Some variations of cognitive attitude views are that emotions have desires or beliefs as necessary components. See Richard Wollheim, *On the Emotions* (New Haven: Yale University Press, 1999); Kenny, *Action, Emotion and the Will*; and Gabriele Taylor, *Pride, Shame, and Guilt: Emotions of Self-Assessment* (Oxford Oxfordshire: Clarendon Press, 1985).

Anger involves a judgment that you or someone you care about has been slighted, usually by an agent who bears responsibility for perpetrating the slight.

The cognitivist attitude view suggests that emotions are or involve attitudes that can be involved in attitude conflicts. After all, to hold a judgment is to commit oneself, minimally, to refrain from having other attitudes that contradict it.[37] Accordingly, if an emotion/judgment conflict arises, it's in violation of a norm of rationality.

Martha Nussbaum, a defender of the cognitivist attitude view, contends that it's not uncommon for an ordinary agent to exhibit this sort of irrationality. She argues that assent is not always deliberately performed, but rather "...habit, attachment, and the sheer weight of events may frequently extract assent from us."[38] This suggests that attitude formation is sometimes passive.

Bennett Helm suggests that emotions are a particularly passive kind of judgment. He observes that when we experience emotions, "the import of our situation impress[es] itself upon us."[39] He describes the sort of judgments that emotions involve as "disclosive." These judgments are revealed by the assortment of emotional reactions we have in various situations. If you were jealous of your lover because you thought that there was a threat to your relationship and you are relieved when your apparent rival moves away, your emotions reveal your commitments: that there was a threat to your relationship and that it's now gone. Emotion/judgment conflicts arise, according to

---

[37] Robert C. Solomon, *Not Passion's Slave: Emotions and Choice* (Oxford: Oxford University Press, 2003): 38; Martha C. Nussbaum, *Upheavals of Thought: The Intelligence of Emotions* (Cambridge: Cambridge University Press, 2001): 36.
[38] Ibid., 38.
[39] Bennett W. Helm, *Emotional Reason: Deliberation, Motivation, and the Nature of Value* (Cambridge: Cambridge University Press, 2001) 66.

Helm, because we have greater control over our judgments than our emotions. Say you judged that there was no real threat of defection from your partner in a relationship but you were jealous nonetheless. When your perceived rival moved away, you still felt relief. On Helm's view, this shows that you assented to the notion that there was a threat, albeit passively, even though you judged otherwise.

Robert Solomon resists the notion that emotions are truly passive, urging us to think of emotions as actions that we undertake. He goes so far as to call emotions "choices" in his early writing.[40] Though some of our emotions might be exceptions—Solomon suggests sudden panic and rage—many of our emotions arguably involve agency in a meaningful sense. Emotions, he argues, involve a history—a history of actions. While they may be initiated without a conscious judgment, he compares the judgments emotions involve in these cases to automatic, habitual, and kinesthetic judgments. They're cultivated, and though they may be unconscious or unreflective, they can be intentional and voluntary, and we can monitor and regulate them.[41] Though Solomon does not discuss the possibility of emotion/judgment conflict explicitly, his view of emotions appears to suggest a greater scope for blame with respect to our emotions in cases of emotion/judgment conflicts than views that place more emphasis on passivity.

There may be a continuum of rational responsibility for emotion/judgment conflicts on the attitude conflict view. If this is the case, we may be more or less culpable for our emotion/judgment conflicts depending on the kind of agency that is

[40] Solomon, *Not Passion's Slaves*, 204.
[41] Ibid., 211-213.

possible with respect to a particular emotion. We may bear stronger culpability for emotions that we are responsible for shaping and maintaining than emotions that we have little agency over. The same might be said with respect to a recalcitrant belief or judgment.

ii. Cognitivist non-attitude views

The cognitivist non-attitude view[42] holds that emotions are not attitudes. Emotions are taken to be representational, but they are not commitment-oriented mental states. On this view, emotions don't imply rational commitments. They have representational content, but it's possible to withhold assent to the representational content even as it impinges upon your consciousness. In this respect, experiencing an emotion/judgment conflict is like experiencing an optical illusion. When you observe a stick that has been partly submerged in a clear glass of water, you can't help but see the stick as though it were bent. What you see does not, however, determine what your attitudes are toward your perception's representational content. Depending on how much you know about light refraction or optical illusions, you might judge that it's bent, that it's not bent, or suspend judgment on the basis of the suspicion that is aroused when you see what happens when you take the stick out of the water and submerge it again. Our emotions, on this view, can persist in much the same way an optical illusion does. You have the option to judge that the emotion's representation is true or false, or to withhold judgment. When you're experiencing an emotion/judgment conflict, you've

[42] These views are often referred to as quasi-judgmentalist views.

judged that the emotion's representational content is false. Thus, you've withdrawn your assent to the emotion's representation.[43]

Different proponents of the cognitivist non-attitude view have compared emotions to a number of representational, but non-commitment oriented mental states such as thoughts, concern-based construals, or inclinations.[44] These are states that can conceivably persist though one has refused or withdrawn assent to any corresponding judgments. You can, for instance, judge that your airplane is not in imminent danger of crashing, but due to the turbulence you're experiencing, vividly maintain the thought that it is.

Proponents of these views choose which mental states to assimilate emotions to on the basis of which components of an emotional episode they want to emphasize. *Thought* and *construal* capture the representational and cognitive style aspects of emotions. Our emotions, involve characteristic ways of attending to objects or situations, and they facilitate the deployment of emotion-congruent concepts in terms of which we organize and understand what we attend to. The *concern* portion of a concern-based construal captures the connection between the representational aspect of an emotion and the feeling and action tendency components. Because emotions are responses to matters of importance to us, our responses are attended by sensations of either a positive or negative quality depending on the whether we deem changes in the

[43] Roberts, *Emotions*, 92; Goldie, "Getting Feelings into Emotional Experience in the Right Way," 237.
[44] Greenspan suggests thoughts. See Patricia S. Greenspan, *Emotions and Reasons: An Inquiry into Emotional Justification* (New York: Routledge, 1988); Roberts suggests concern-based construals. See Roberts, *Emotions*; Brady suggests inclinations. See Michael S. Brady, "The Irrationality of Recalcitrant Emotions," *Philosophical Studies* 145.3 (2009).

status of these matters to be good or bad for us. They also dispose us to take appropriate actions according to what the situation demands so that matters of concern to us are adequately attended. *Inclinations* capture the action tendency and cognitive style components. Our emotions incline us to devote whatever cognitive resources and physical energy we can afford to matters of importance to us, especially if the matter appears urgent. Urgency is reflected in the intensity of the emotion and its cognitive and practical effects.

Since none of these mental states—thoughts, concern-based construals, or inclinations—necessarily implies assent, there is no danger of an attitude conflict if they persist in the face of a contrary judgment. The accounts of the irrationality of emotion/judgment conflicts offered by such views do not, therefore, describe them as conflicts between attitudes. Instead, they argue that such conflicts violate norms of rationality other than the ones that prohibit outright attitude conflicts.

Robert Roberts suggests that emotions cause irrationality because, being closely connected to our personal concerns, their motive force is not always lessened by our judgment that their representational content is false. They thus influence us to do things that we judge we shouldn't.[45]

In addition to their influence over our actions, Michael Brady argues that, because emotions have effects on our attention and concept deployment, emotions prime us to form beliefs and judgments contrary to what we've judged to be correct. To be inclined to form attitudes that we judge we shouldn't, according to Brady, is to

[45] Roberts, *Emotions*, 87-9.

violate norms of rationality associated with the notion that we ought to be guided by our judgment faculty.[46]

iii. Neo-Jamesian views

For views rooted in the Jamesian tradition, the essential components of an emotion are bodily change and bodily feeling. On James's view, there can be no emotion without feeling. In defense of this view, he offered a thought experiment which asks us to imaginatively abstract feeling from an emotional episode and see what's left. He writes,

> Can one fancy the state of rage and picture no ebullition of it in the chest, no flushing of the face, no dilation of the nostrils, no clenching of the teeth, no impulse to vigorous action, but in their stead limp muscles, calm breathing, and a placid face? The present writer, for one, certainly cannot. The rage is as completely evaporated as the sensation of its so-called manifestations, and the only thing that can possibly be supposed to take its place is some cold-blooded and dispassionate judicial sentences, confined entirely to the intellectual realm, to the effect that a certain person or persons merit chastisement for their sins.[47]

He concludes that "emotion dissociated from all feeling is inconceivable."[48] Since the feelings we experience during an emotional episode are of changes in the body, emotions are felt experiences of changes in the body.[49]

The success of this thought experiment is, however, doubtful. It may very well be possible to conceive of an emotion that is unfelt. Some cognitive theorists argue that either no feelings or no particular feelings are necessary components of an emotion.[50]

---

[46] Brady, "The Irrationality of Recalcitrant Emotions," 428.
[47] James, "What is an Emotion?" 132.
[48] Ibid.
[49] Ibid., 132-133.
[50] Nussbaum, *Upheavals of Thought*, 29-30; Roberts, *Emotions*, 66; Donald Davidson, "Hume's Cognitive Theory of Pride," *The Journal of Philosophy* 73.19 (1976): 745.

Even neo-Jameseans have suggested that the processes underlying emotional feelings can run their courses unconsciously. Damasio's neuroimaging studies show that areas of the brain responsible for tracking changes in the body activate during an emotional episode whether we're conscious of those changes or not. On the basis of this research, Jesse Prinz concludes that the bodily change component of an emotion can go on unfelt.[51]

Prinz, who defends the neo-Jamesian view that emotions are essentially perceptions of the body, makes an important addition to the view: an account of emotional representation that is compatible with non-cognitivism. An emotion's representational content is usually identified with a cognitive emotion trigger and its semantic content. On Prinz's view, emotions have a variety of triggers including cognitions, but he argues that some emotion triggers are non-cognitive.

These apparently non-cognitive triggers are simple perceptions originating in the thalamus. The thalamus registers only course visual features. LeDoux hypothesized that there is a direct pathway from the thalamus to the amygdala, where emotions are initiated. Emotions are thus triggered, in at least some cases, without mediation from the neocortex, which is responsible for the recognition of perceptual inputs.[52] According to Prinz, nothing about the emotions resulting from these sorts of triggers deserves the label "cognitive." For something to be cognitive, he argues, it has to be a representation

[51] Prinz, *Gut Reactions*, 58, 201-205.
[52] Joseph E. LeDoux developed a "dual pathway model" of emotion. The thalamus sends signals to both the neocortex and the amygdala, but the emotion is initiated before recognition is achieved in the neocortex and the amygdala, but the emotion is initiated before recognition is achieved in the neocortex and the amygdala, but the emotion is initiated before recognition is achieved in the neocortex. See Prinz's interpretation and discussion of LeDoux's view: Ibid., 34.

over which we exercise executive control. That is, we're able to control it directly by activating it, maintaining it, or manipulating it in working memory. Simple perceptions are passive states with exogenous causes, and on Prinz's view, emotions are perceptions of the body whose content is purely non-conceptual. Any cognitive effects of an emotion are not a part of the emotion itself.[53]

Prinz offers an alternative account of emotional representation designed to obviate the need to appeal to cognition to explain how perceptions of the body could represent matters of significance outside of the body as dangerous, fearsome, happy, and so forth. Emotions are, on his view, "embodied appraisals." They are states of the body that track "core relational themes."[54] The appraisal theorist Richard Lazarus (discussed above) coined the term core relational theme to designate a summary appraisal that ties together a number of component appraisals typical of a particular kind of emotion. The core relational theme of fear, for example, is danger, and the core relational theme of guilt is self-blame.[55] Prinz retools the notion of a core-relation theme to fit his purposes. Instead of a summary of cognitive appraisals, it is a category of emotional stimuli that reliably co-occur with a particular set of bodily changes.[56] The bodily changes associated with fear co-occur with stimuli that can be categorized in terms of danger, and the bodily changes associated with guilt co-occur with stimuli that can be categorized in terms of self-blame, and so forth.

---

[53] Ibid., 38-38.
[54] Ibid., 64-66.
[55] Smith and Lazarus, "Appraisal Components, Core Relational Themes, and the Emotions," 238.
[56] Prinz, *Gut Reactions*, 53-55, 67-69.

Borrowing Dretske's account of mental representation, Prinz argues that emotions *represent* core relational themes. On Dretske's view, a mental state can represent that which it reliably detects.[57] Because our emotional responses co-occur with core relational themes, and because they appear to have been selected for by evolutionary processes that promote advantageous behavior in specific types of situations, emotions might reasonably be said to detect core relational themes.[58] Bodily changes associated with fear detect dangerous situations, so they represent a situation as a dangerous one, and bodily changes associated with guilt detect situations which call for self-blame, so they represent a situation as one in which one is blameworthy.[59]

Though you might expect proponents of non-cognitivist neo-Jamesian views to hold that emotions are *arational* states that are not directly implicated in any irrationality they might cause, this conclusion isn't taken for granted by the view's contemporary proponents. Like the defenders of the cognitivist non-attitude view, proponents of non-cognitive neo-Jamesian views acknowledge that there is an apparently significant difference between an errant perception and an errant emotion that makes normative assessment possible. That is, we do exercise some agency in the

[57] Dretske's view is that there are two criteria for being representational: representations carry information, and they can be in error. A minimal condition for carrying information is reliable co-occurrence. Prinz argues that emotions meet these criteria. The bodily changes we perceive when we're experiencing emotions are caused by the detection of situations that fit the profiles of core relational themes. Thus, the first condition is met. The second condition is met because emotions can be triggered in error, resulting in a representation of a core relational theme that doesn't fit the actual situation. Ibid., 53.
[58] Ibid., 53-55, 67-69.
[59] See Karen Jones, "Quick and Smart? Modularity and the Pro-Emotion Consensus," *Canadian Journal of Philosophy* 36.32 (2006): 19-25 for helpful summary of the Prinz's view and some incisive criticisms of it.

regulation of our emotions. Emotions can be calibrated via therapy or cognitive training to produce better or more accurate outputs. Insofar as it is in our power to control our emotions, preventing inapt emotions from occurring, we bear responsibility for them. (Cases of phobias, affective disorders, chemically induced affective states, and any other emotions that we don't have any meaningful control over are exceptions.)[60] Tappolet argues that this is enough for norms of rationality to apply to emotions generally.[61]

Prinz argues that there is a sense in which emotions are arational, but another sense in which the familiar categories of rational assessment that we normally reserve for the assessment of attitudes extend to emotions. He explains by drawing an analogy between emotions and arguments. The mental representations that elicit emotions are like premises. An emotion's triggers, which are regulated by what Prinz calls a "calibration file," are like inference rules. Calibration files are stored memories or representations that trigger a particular type of emotion. They resemble inference rules because they connect representations to emotions. So, for example, if the appearance of a snake is a trigger for fear, we can compare this connection to the inference rule "*If you see a snake, experience fear*."[62] We are responsible for the "premises" and "inferences" that result in a "conclusion" (i.e. and emotion). But according to Prinz, emotions are at a level of remove from the initiation pathways that they are "inferred" from. Once a calibration file has been established, it's not strictly speaking irrational for you to

[60] Prinz, *Gut Reactions*, 237-8.
[61] Christine Tappolet, "Emotions, Perceptions, and Illusions," in *The Crooked Oar, the Moon's Size and the Kanizsa Triangle: Essays on Perceptual Illusion*, edited by Clotilde Calabi (MIT Press, forthcoming) 33.
[62] Prinz, *Gut Reactions*, 100, 237-8.

experience an emotion when you encounter situations or objects that contain those triggers. We might, however, be blameworthy for any irrational past actions and attitudes that shaped the emotion's calibration file. Prinz concludes that emotions are a special case of rationally assessable attitudes due to the mixture of active and passive dimensions that they possess. He explains,

> Emotions tend to be more amenable to normative assessment than paradigm instances of perception, because their elicitation is more dependent on prior mental states and rules. Even so, there is always a sense in which emotional responses are beyond our control. There is a sense in which the most heinous passion is as innocent as seeing an afterimage.[63]

4. Discussion: Emotions, attitudes, and cognition

The issue of whether the cognitive appraisal component of an emotion is an essential part of emotional experience marks a pivotal divide in both psychological theories and philosophical views of emotion. Both appraisal and two-component theories posit that emotions have cognitive appraisals as essential components. For appraisal theories, they're at the beginning of the sequence of the emotional episode. For two-component theorists, they come after or in tandem with the bodily change component, and the two components combine to give an emotion its particular quality. Philosophical cognitive views of emotion also regard the appraisal component of an emotion as necessary. They divide into views on which appraisals are judgments and views on which appraisals are non-commitment-oriented mental states such as evaluative thoughts or construals. The opponents of these views emphasize the bodily change and feeling components of an emotion, regarding the cognitive components of

[63] Prinz, *Gut Reactions*, 240.

an emotion as non-essential. Philosophical non-cognitivist views reject the notion that emotions could be judgments or any other variety of cognitive mental state.

The controversy over the relationship between emotions and cognitive appraisals came to a head in the 1980s after Robert Zajonc published his now classic (1980) paper "Feeling and Thinking: Preferences Need No Inferences."[64] Citing research on exposure effects, he argues in the paper that emotions can occur with simple perceptual inputs. Exposure effects studies show that subjects who have been exposed to a stimulus repeatedly, albeit only briefly, tend to demonstrate increased liking of the stimulus. The effect holds even though these subjects are unable to identify the stimulus as something that they've encountered before. As a result of these findings, Zajonc concluded that at least some of our emotions are generated without cognitive input. Mere exposure to a stimulus, without cognitions regarding its identification or significance, were apparently sufficient to generate an emotional state directed toward that stimulus when it was subsequently presented. Thus Zajonc infers that cognitive appraisals are not an essential component of an emotion.

Zajonc's argument was highly controversial and divisive. As a number of critics pointed out, the findings of the exposure study are consistent with the notion that whatever emotions arise in participants in Zajonc's studies are preceded by an unconscious appraisal.[65] In response to the criticism that his paper generated, Zajonc

---

[64] Robert B. Zajonc, "Feeling and Thinking: Preferences Need No Inferences," *American Psychologist* 35.2 (1980).

[65] See Richard S. Lazarus, "Thoughts on the Relations between Emotion and Cognition" *American Psychologist* 37.9 (1982): 1022; Jan De Houwer and Dirk Hermans, "Do Feelings Have a Mind of Their Own?" in *Cognition and Emotion: Reviews of Current Research and Theories*, edited by Jan De Houwer and Dirk Hermans (New York: Psychology Press, 2010) 39-40.

wrote a companion paper (1984) that clarified and extended his arguments. Most importantly, Zajonc proposed a definition of "cognition" as necessarily involving "transformation of a present or past sensory input," or "mental work."[66] By this measure, he claims, states triggered by untransformed sensory inputs are non-cognitive.

Putting controversies about the hypothesis that simple perceptions can serve as emotion triggers aside,[67] if we accept Zajonc's definition of cognition, the claim that

---

[66] Robert B. Zajonc, "On the Primacy of Affect," *American Psychologist* 39.2 (1984): 118.

[67] LeDoux's research on fear is frequently cited as decisive support for Zajonc's claims. LeDoux showed that, in rats, a fear response could occur with activity in subcortex, but not the neocortex. On the basis of these results, he conjectured that a fear response could occur in humans via a subcortical route to the amygdala too. The subcortex takes low level perceptions as inputs, and these inputs do not undergo processing until they reach the neocortex. If a low level perception can skip neocortical processing and go directly to the amygdala by a subcortical route, then no "mental work," as Zajonc calls it, is necessary for a fear response to be initiated. See Joseph E. LeDoux, *The Emotional Brain* (New York: Simon and Schuster, 1996); Joseph E. LeDoux, "Cognitive-Emotional Interactions in the Brain," *Cognition and Emotion* 3.4 (2008); and discussion in Prinz, *Gut Feelings*, 34. Though LeDoux's research is widely cited, it has generated a significant amount of controversy in its own right. Questions have been raised about the anatomical dissimilarities between rats and humans, and the failure of consistent alignment of human and rat responses to emotional stimuli in subsequent research. See Richard J. Davidson, Andrew Fox, and Ned H. Kalin, "Neural Bases of Emotion Regulation in Nonhumam Primates and Humans," in *Handbook of Emotion Regulation*, edited by James J. Gross (New York: Guilford Press, 2007) 47-48; Kevin N. Ochsner and James J. Gross, "The Neural Architecture of Emotion Regulation," in *Handbook of Emotion Regulation*, edited by James J. Gross (New York: Guilford Press, 2007) 88-89. It's frequently pointed out that, in humans, emotions involve multiple areas of the brain in addition to the subcortex, and that many emotion triggers are higher-level perceptions or cognitions. The controversy is surveyed in Rick Anthony Furtack, "Emotion, the Body, and the Cognitive," *Philosophical Explorations* 13.1 (2010): 53. Note also that in his more recent writings, LeDoux himself denies that amygdala activation alone deserves to be called "fear" even though he uses such language in his early work. An emotion such as fear involves a threat detection system—the amygdala is at least sometimes, though not always, involved in threat detection. It also involves a conscious experience which isn't fully reducible to the activities of the threat detection system. The fear response involves subsequent activation of body and brain states. Both of these may be tied to other survival circuits that don't involve the amygdala, and they might be heavily shaped by higher cognitive functions involved in the labeling, assessing, and

emotions can occur without cognitions may very well follow. This would be a devastating blow to cognitivist philosophical views of emotion. It's the notion that emotions are necessarily accompanied or caused by cognitions that allows cognitivist to compare emotions with judgments, thoughts, and other cognitive phenomena. However, Zajonc's definition is just as controversial as his conclusion.[68] First and foremost among his critics is Richard Lazarus, the appraisal theorist. On Lazarus's view, our cognitive and emotional processes are thoroughly entwined. When we respond to emotional stimuli automatically, our beliefs, past appraisals, motives, expectations, and patterns of attention to matters of importance are brought to bear—unconsciously—in the formation of present appraisals.[69] What Zajonc calls the "minimal cognitive engagement" involved in emotions that are supposedly generated sans cognition is, according to Lazarus, sufficient for emotions to count as cognitive. The cognitions in these cases may be simple and automatic rather than complex and conscious, but they are nevertheless cognitions on Lazarus's view.[70]

Both Zajonc and Lazarus have been criticized for their narrow focus on competing definitions of cognition. The debate pits definitions with narrower extensions against definitions with broader ones aimed at excluding or including a limited subset of

---

regulation of emotions. See LeDoux, "Rethinking the Emotional Brain," 665-666; Joseph E. LeDoux, "The Slippery Slope of Fear," *Trends in Cognitive Sciences* 17.4 (2013): 156; and Joseph E. LeDoux, "What's Basic about the Brain Mechanisms of Emotion?" *Emotion Review* 6.4 (2014): 319.

[68] Bernard J. Baars, "Cognitive Versus Inference," *American Psychologist* (1981): 223; Leventhal and Scherer, "The Relationship of Emotion to Cognition," 4.

[69] He explains, "Information processing as an exclusive model of cognition is insufficiently concerned with the person as a source of meaning." Lazarus, "Thoughts on the Relationship between Emotion and Cognition," 1020.

[70] Ibid., 1022.

mental activities from the category of cognitions. It's alleged that the debate's narrow focus on these categories and their applicability to an emotion's early stages may prevent researchers from appreciating the interdependence of affect and cognition in a dynamic process that spans across the whole emotional episode.[71]

In spite of the criticisms the debate about emotions and cognition has generated, and the growing sense that the debate has become a distraction from more tractable theoretical inquiries, philosophical theories developed in recent years have largely been constructed with the existing boundaries of the cognitivist and non-cognitivist divide in mind.[72] Cognitivist attitude and non-attitude views appeal to the work of cognitivist researchers, and non-cognitivist views are supported by the research from the other (non-cognitivist) party-line. Here, I will make no definite pronouncements or predictions regarding the debate's outcomes, present or future. I propose that, instead, having learned what we can from the debate at present, we consider a view that has little staked on its outcome: the attitude view.

i. The attitude view

The attitude view of emotions that I develop here is, in a sense, hardly new or novel. On the contrary, it's often presumed without argument by philosophers who are not participants in the particular debates in philosophy of emotion addressed in this

[71] Andrew Ortony, Gerald L. Clore, and Allan Collins, *The Cognitive Structure of Emotions* (Cambridge: Cambridge Univ) 5; Leventhal and Scherer, "The Relationship of Emotion to Cognition," 4-8.
[72] See's review in Furtack, "Emotion, the Body, and the Cognitive," 52, 58.

dissertation.[73] The lack of appropriate mooring has, however, made it an easy target for critics. In this dissertation, I have undertaken to give it a substantial defense, and to show that it has merit that should not be overlooked or underestimated. A related attitude view has also been proposed and defended by Deonna and Terroni.[74] They have developed their view with an eye toward resolving difficulties regarding the identification of an emotion's content. The attitude view that I defend addresses a different set of difficulties, which stem from emotion/judgment conflicts and rationality concerns.

The core features of the attitude view I defend are as follows. First, emotions are not subsumed by our other existing attitude concepts (e.g. judgments, beliefs, and so forth); they are a unique category of attitudes. Although emotions and judgments both involve evaluation and assent, the attitude view does not depend on there being any further similarities between emotions and judgments. Second, our attitudes, including our emotions, have state manifestations and a supporting dispositional edifice. Third, like attitudes generally, our emotions are typically reason responding. I elaborate on the second and third features of the view in the remainder of this section.

Attitudes—beliefs, judgments, desires, and intentions are paradigm instances—are not transitory cognitive actions or states. They are sustained by underlying

[73] It should be noted that there is another often assumed view corresponding to the notion that emotions are blind bodily feelings.

[74] Julien A. Deonna and Fabrice Teroni, "In What Sense are Emotions Evaluations?" in *Emotion and Value,* edited by Sabine Roeser and Cain Todd (Oxford: Oxford University Press, 2014); and Julien A. Deonna and Fabrice Teroni, "Emotions as Attitudes," *Dialectica* 69.3 (2015). I believe that my approach is generally compatible with theirs as it's developed in these sources though my view is adapted to a different set of issues in the philosophy of emotion regarding emotions and rationality.

dispositional structures in long-term memory that result in particular state-manifestations. Emotions are no different. To undergo the multicomponential recursive processes that make up an emotional episode is to be in an emotional state. To have a stable tendency to be in certain states with respect to a certain stimulus or class of stimuli is an emotional disposition. For example, if I'm presently experiencing joy because I've just found out that my paper has been accepted for publication, I'm undergoing a process that involves a particular kind of appraisal, a particular cognitive style, bodily changes, and so forth. I'm thus currently in a state of joy. Hours later, when I'm no longer undergoing the same emotional process, I continue to be disposed to re-enter a state of joy if I am reminded of the circumstance that elicited the emotion. I'm also disposed to experience similar emotions regarding successes which carry a similar personal significance for me, particularly other publications. I thus have a disposition toward the circumstance that produced my original emotional state and a disposition to experience similar emotions in relevantly similar circumstances.[75]

Emotional dispositions consist in the structures in long-term memory that shape our emotional states.[76] Research on emotional development suggests that the schema of memories and concepts that form our emotional dispositions beyond infancy grow out of a set of innate stimulus responses we were born with. We develop associations

[75] Wollheim proposes a similar state/disposition distinction. According to his view, mental states as episodic and transient. Dispositions, by contrast, are rooted in persisting modifications that underlie sequences of mental states, so they persist beyond the duration of particular state manifestations. Mental states are accessible to subjective experience while dispositions are only experienced when they manifest in states. See Wollheim, *On the Emotions*, 1-2.
[76] Ibid.; See also Peter Goldie "Wollheim on Emotion and Imagination," *Philosophical Studies* 127 (2006): 1-3.

between these innate responses and the sorts of situations that trigger them over time, and our innate responses are transformed by learning as we respond to new information and novel situations. From our emotional experiences, we develop schemas of emotion concepts, associations, and memories that are coupled with our goals, motivations, and expectations.[77] These schemas determine what sorts of perceptions or cognitive appraisals trigger our emotional states. They also overlay our emotional experiences with the sorts of cognitive elaborations they tend to engender, which in turn steers an emotion's course as the situation unfolds.

Here, I've suggested that emotions are attitudes that admit of both episodic states and persisting dispositions, but it's important to note that not all dispositions of mind are attitudes. Our dispositions to feel hot and cold or to perceive a stick submerged in water as though it were bent are not attitudes because they do not arise as responses to considerations in favor of such states. Any "reasons" that we can offer for being in such states are merely explanations of their causes, which need not provide normative support one's being in such a state.[78] For something to be an attitude, it must be the sort of state that is generally responsive to reasons that count in favor of adopting, cultivating, or sustaining it. By the same token, it must be the sort of state that we can generally revise in light of new or different reasons.

---

[77] See Moors, "Theories of Emotion Causation," 18; Leventhal and Scherer, "The Relationship of Emotion to Cognition," 9.

[78] Here, I have in mind a distinction between explanatory reasons and reasons in the standard normative sense. See T. M. Scanlon, *What We Owe to Each Other* (Cambridge: Belknap Press, 1998) 21. This is similar to the notion of a reasons-for-which explanation in Kate Nolfi, "Which Mental States are Rationally Evaluable, and Why?" *Philosophical Issues* 24 (2015): 43-44; and the notion of question-settling reasons in Pamela Hieronymi, "Controlling Attitudes," *Pacific Philosophical Quarterly* 87 (2006): 51.

Our ability to revise an attitude in light of reasons need not be perfect or immediate. Because our attitudes are dispositions to respond to reasons, we don't manipulate, control, or perform mental work on them directly. We wield influence over our attitudes by engaging in various activities of reflection. These are cognitive mental actions such as thinking, attending, reasoning, inferring, imagining, and weighing that we utilize when we're considering our reasons to judge, believe, desire, or intend. Thoughts, images, and recollected concepts are the states that we manipulate and perform mental work on.[79] The attitudes that emerge from these activities have an importantly passive dimension. We are constrained by the reasons that we are confronted with,[80] and we are very often compelled by the force of history and habit in responding to our reasons.[81] Some of our attitudes have deep roots extending through related memories, attitudes, and mental routines. Attitude change in such cases may

[79] Judging is sometimes treated as mental action. See for instance, Christopher Peacocke, *Being Known* (New York: Oxford University Press, 1999) 19-20. Others argue that it is a non-voluntary attitude that arises spontaneously and nonvoluntarily. See for instance Galen Strawson, "Mental Ballistics or the Involuntariness of Spontaneity," *Proceedings of the Aristotelian Society* 103 (2003): 237. Strawson denies that judgment, thought, or reasoning count as mental actions. In fact, very few things do, according to his analysis. We may perform actions that initiate mental states, but most of what occurs is "mental ballistics" that occurs beyond the scope of what's voluntary.

[80] See Williams's influential arguments against belief voluntarism. Bernard Williams, "Deciding to Believe," in Problems of the Self*: Philosophical Papers 1956-1972*, (Cambridge: Cambridge University Press, 1973) 136-151.

[81] See Connor McHugh, "Epistemic Responsibility and Doxastic Agency," *Philosophical Issues* 23 (2013): 133-135; Josepha Toribio, "What We Do When We Judge," *Dialectica* 65.3 (2011): 351; Eric Schwitzgebel "Acting Contrary to Our Professed Beliefs or the Gulf Between Occurrent Judgment and Dispositional Belief" *Pacific Philosophical Quarterly* 91 (2010): 535; Nolfi, "Which Mental States are Rationally Evaluable," 54-56; Strawson, "Mental Ballistics or the Involuntariness of Spontaneity" 237.

require incremental efforts of reflection and self-regulation over an extended period of time.

Emotions are, I contend, generally reason responding in the manner of an attitude. Since emotions are responses to matters of value for an individual, they are responses to both truth- and value-relevant reasons. To fear some object or state of affairs is to regard the situation as (truly) a threat of some kind, and to regard it as the antithesis of something valued—one's safety and well-being. When new or different considerations regarding these matters are brought to bear on our emotional attitudes, attitude change typically occurs. If on reflection, it becomes apparent that what once appeared to be a threat is not truly dangerous, fear will normally be extinguished. Our deeply rooted emotional dispositions may be difficult to change, but even then, change can ordinarily be achieved with reflection and self-regulation, at least incrementally.[82] There are, of course, exceptional cases in which emotions are recalcitrant, and thus fail entirely to be reason responding. Nevertheless, we are generally able to shape our emotional states and dispositions with reflection on our reasons.

The attitude view has some important advantages over the alternatives. First, as I alluded to at the end of the previous section, it commits us to only a minimal cognitivist thesis that doesn't hinge on whether it's true that all emotions are preceded by cognitions. Second, it provides us with the most plausible account of emotion/judgment conflicts. It explains how and why they are irrational and how they ought to be

[82] I argue that emotions are reason-responding in chapter 3. I borrow from the existing research on emotion-regulation to show that the activities of reflection can be used to influence our emotional states and dispositions to achieve reason-oriented changes in our emotions.

adjudicated in a way that respects our reasons. I elaborate on these claims in 4.ii and iii. Section 4.iv also provides a preliminary sketch of an account of emotion/judgment conflicts and its implications according to the attitude view.

ii. Emotions, dispositions, and cognition

Whether emotions are initiated by cognitive appraisals in the narrow sense implied by Zajonc's or Prinz's definitions is of little consequence for the attitude view of emotions.[83] Attitude *states* may involve cognitions coupled with assent or motive states, but as dispositions to respond to reasons, they need not be constituted by present cognitions. The attitude view, accordingly, only commits us to a minimal cognitivist thesis according to which cognitions are involved in various parts of an emotion's trajectory, with our earliest infantile emotions as the only likely exceptions. Cognitions generally play an important role in the formation of our emotional dispositions, and they are also typically involved in our emotional states as the emotional episode unfolds over time. Beyond infancy, cognition will usually play some role, however small, in the initiation of an emotional state because we are always processing information about our circumstances and these cognitions have an impact on our expectations and our subsequent emotional responses. Even if there are no cognitions involved in the initiation of some emotions, for an adult human being who possesses emotional

[83] Zajonc proposed that cognition requires transformation of a sensory input, or more succinctly, "mental work." See Zajonc, "On the Primacy of Affect," 118. Prinz proposes that something is cognitive if and only if it's a representation over which the organism exercises "top down," or executive control. Roughly, this means that the organism is able to exercise direct control over a representation rather than or in addition to indirect control (e.g. manipulation by shifting attention). See Prinz, *Gut Reactions*, 38-9.

schemas that will become activated once an emotion has been initiated, cognitions will nevertheless play a role in the ensuing emotional state. Cognitions caused by the emotion's initiation will impact the way that the emotion unfolds and develops.

If an emotional state is triggered by apparently non-cognitive perceptions, these triggers are still, as Frijda describes them, "squeezed between much more complex cognitive activities."[84] Long before an emotion is initiated, Frijda points out, you had a lifetime of cognitive events to shape what sorts of perceptions would elicit the particular emotional response that it did. If your emotion is guilt, the guilt you experience will be contoured by "previous doubts, irritations, ruminations, and apprehensions" which, themselves, arose from experiences of norm violations that you've perpetrated or witnessed. After the emotion is initiated, cognitive elaboration occurs. Cognitive elaboration connects information from your emotion schema or script to your perceived context. In guilt, you identify a particular harm or norm violation, you look for offended parties, you consider your intentions in undertaking the actions that you're guilty of, you assess coping potential and the expected efficacy of available responses such as keeping your distance, making amends, and offering apologies. Even as cognitive elaboration is underway, you're monitoring your environment and taking in new information. This new information may, in turn, modify your emotion or elicit a different one in addition to or in place of your present guilt.[85]

It might be protested that, as a social or moral emotion, guilt may necessarily require complex knowledge or concepts that simpler emotions such as fear may not. Say

[84] Frijda, "The Place of Appraisal in Emotion," 382.
[85] Ibid., 383.

your emotion is a startle or a burst of fright at hearing a crackling noise as you walk in a solitary wooded area. Frijda remarks, "…even there one has been walking around with all sorts of expectations against which the crackle is perceived."[86] Like all of our emotions, fear generally arises against a background of goals and expectations that are activated, though perhaps unconsciously, in such situations.[87] You don't expect to be disturbed in a solitary wooded area, for example, but you are aware that there could be unknown dangers in the vicinity, and you want to be safe. These sorts of background attitudes, goals, and expectations dispose you to experience fear on hearing a crackle in the woods. In another situation, the very same noise would elicit a very different response if it elicited your notice at all. Since your emotional state on hearing the crackle in the woods is an expression of your disposition for fear in a specific kind of circumstance, it is an expression of an attitude. This is not to deny that there can be unconditioned emotional responses, but after early stages of human development, there aren't likely to be many emotional responses that aren't shaped by concepts, attitudes, expectations, and goals that are activated in particular contexts either before or after the emotion is initiated.

The attitude view, I conclude, commits us to plausibly modest claims about the role of cognitions in emotion. There is no need to claim that emotional states are necessarily initiated by cognitive appraisals to show that emotions are thoroughly infused with cognitions in the vast majority of cases.

---

[86] Ibid., 382.

[87] For a review of research on automatic and unconscious goal activation, see Gráinne M. Fitzsimons and John A. Bargh, "Automatic Self-Regulation," in *Handbook of Self-Regulation: Research, Theory, and Applications*, edited by Kathleen D. Vohs and Roy F. Baumeister (New York: Guilford Press, 2011) 151-170.

i. Emotion/judgment conflicts: a sketch

In anticipation of the arguments to come in the chapters that follow, here I briefly sketch the view of emotion/judgment conflicts that the attitude view recommends.

Say that you're experiencing guilt, and you can't dispel it though you judge that there were no better, less regrettable alternatives to the actions that you feel guilty about. If there were no better alternative actions available under the circumstances, you reason, you cannot bear the sort of responsibility for your actions that would warrant your guilt. Still, you have a sense of having let others down in a serious manner, and you feel responsible for the harms that resulted from your actions. Your thoughts and motivations are attuned to the need to redress, compensate, or apologize for the harms that your guilt targets.[88] You rehearse the reasons that led you to judge that your guilt isn't warranted by your circumstances, but the emotional episode continues to run its course without successful interference. The attitude view suggests that you're experiencing an attitude conflict. You harbor one attitude—a judgment—that you aren't responsible for the harms your actions caused, so guilt is inappropriate, and you harbor a conflicting attitude—an emotion—of guilt.

Each attitude is a response to a set of considerations. In your guilt, you're responding to features of the situation that support a sense that you're responsible for a harm, and that the harm needs to be redressed in order to restore your conceit to moral goodness and your favorable standing in the community. Guilt represents your situation

[88] Frijda, "The Place of Appraisal in Emotion," 364-5.

as one in which you are blameworthy. Your judgment faculty, on the other hand, is responding to features of the situation that don't fit with your sense of having been personally responsible for any harms you perpetuated. You focus on external forces that were at work in the situation which seem to have made it impossible (or at least impractical) for you to behave otherwise than you did. Your judgment represents your situation as one in which you are not blameworthy.

Attitude conflicts of this sort are intuitively irrational, and rationality requires a resolution of the conflict to restore consistency and coherence. As long as the two representations clash, each has the potential to drive your future thoughts and actions as you navigate the situation that your guilt is directed toward.

A conflict may dissolve easily if the attitudes involved aren't deeply entangled with your most entrenched background attitudes or habits. Imagine that after some further reflection on your reasons for feeling guilty, you resolve the conflict by a change in either your emotion or your judgment. On reflection, you might discover that you judged wrongly and there *were* alternatives available that you should have known about that would have avoided the harms your actions caused. Or, you might discover that, with your emotion, you were reading responsibility into the situation although you really *weren't* culpable for your actions under the circumstances. There were no realistic alternatives available to you. The circumstances were thus importantly different from previous situations that appropriately elicited your guilt, so your disposition for guilt simply wasn't attuned to the sorts of reasons that it ought to have responded to in this case. Emotion/judgment conflicts can usually be resolved with reflection in ways I discuss at length in chapter 3. Our emotions are generally sensitive to reasons regarding

their correctness and appropriateness. After some more reflection on the reasons that (rightly) suggest that guilt is inappropriate in this circumstance, guilt should dissipate.

The difficulty of resolving emotion/judgment conflicts increases to the extent that the attitudes involved issue from stable dispositions that may take a concerted, long-term effort to change. Perhaps you have a deeply rooted habit of making judgments that deflect guilt because you are afraid to take responsibility for your actions and it causes you to avoid or suppress evidence of your culpability. Or perhaps you have a long-standing habit of experiencing guilt for matters over which you have no real or substantial control. Your dispositions for guilt are thus acutely oversensitive to reasons that could indicate some kind of responsibility. In the worst cases, we have attitudes that are altogether recalcitrant. Recalcitrant attitudes have become insulated from the influence of reflection. If no considerations bearing on your control over a situation could move your judgment, it is recalcitrant, and if no considerations bearing on your powerlessness in a situation could allay your guilt, it is recalcitrant.

Insofar as some degree of agential control or influence is necessary in order to assign responsibility for an attitude, it may not be appropriate to extend blame to persons whose attitudes are wholly irresponsive to reflection on their reasons.[89] Still, most of the time, we are able to guide and shape our attitudes with reflection. To the extent that we are capable of doing so, we are responsible for our emotions, our judgments, and the consistency and quality of our attitudes generally.

[89] See Nolfi, "Which Mental States are Rationally Evaluable," 44-47; and McHugh "Epistemic Responsibility and Doxastic Agency," 142-144 regarding the relationship between agential control and rational responsibility.

iii. Looking Forward

As I develop this view of emotion and emotion/judgment conflict in the rest of the dissertation, the advantages of the view over alternatives will become apparent. The most important of these is that alternative views of emotion/judgment conflicts and the irrationality they're implicated in can't do justice to our emotional faculties as reason-responding faculties. Our emotions bring our history, our habits, and our background attitudes to bear on matters of personal importance to us in the present. When our emotional faculties misfire because our past experiences aren't suited to a correct view of the present, our emotions generally respond to reflection on differences between past emotional experiences and the aspects of our current circumstances that are truly different or novel, just as any attitude that is shaped or informed by past experience should.

Since emotions are reason-responding faculties, their conflicts with other reason-responding faculties ought to be treated as attitude conflicts. As attitude conflicts, emotion/judgment conflicts are irrational in virtue of the conflict, and rationality demands conflict resolution in accordance with our reasons. Which attitude should prevail is generally indeterminate until further reflection reveals which attitude, if either, represents an apt response to a reason-set that is appropriately inclusive and with reasons properly weighted. The view does not proffer presumptive favor to our judgment faculty except in circumstances in which it is quite clear that one's judgment faculty is well-suited to detecting and accurately assessing all of the relevant reasons regarding the matter at hand and it is equally clear that one's emotion faculties are seriously disordered or untrustworthy.

Other views cannot afford the same sort of standing to emotions that are in conflict with our judgments. On other views, emotions typically lack the rational authority to seriously challenge the deliverances of our judgment faculty.

If emotions are cognitive non-attitudes such as thoughts or construals, they don't involve reason-based commitments. Defenders of these views claim that emotions that conflict with our judgments are irrational due to their unreasoned persistence or their interference with our judgment faculty's guidance of thought and action. On these views, emotions are the sole source or cause of any irrationality they're implicated in. This leaves little room for exceptions in cases in which our emotional states or dispositions are better founded than the deliverances of our judgment faculty.

On the view that emotions are non-cognitive and therefore arational, they can at best be treated as defeasible evidence of reasons in favor of revising one's judgment. There is no obvious or necessary normative pressure, then, to resolve the conflict. The only clear requirement the view entails with respect to the judgment in an emotion/judgment conflict is to incorporate whatever evidence the emotion represents into the set of evidence that supports maintaining the judgment. If there is reason to revise the emotion, it's to prevent it from potentially becoming disruptive. These sorts of requirements and incentives don't do justice to either our emotions or our judgments as a reason-responding, but fallible states. Judgments are given disproportionate deference and emotions too slight regard.

Responding to our emotions rationally requires giving all of our faculties their due consideration. Judgments can be short-sighted, ill-founded, or mistaken responses to reasons, just as our emotions can. Conversely, our emotions and our judgments can

both be well-regulated and carefully monitored, and they can be founded on careful reflection. If an emotional faculty is of the latter sort, it isn't obviously the cause or source of the irrationality of an emotion/judgment conflict. An adequate view of emotion/judgment conflicts ought to make room for the possibility that the judgment is the cause or source of the irrationality. The attitude view does just that.

Chapter 3

Emotions and Reasons

By measures of frequency, centrality, and emphasis in the philosophical literature on attitudes, it is apparent that the distinguishing feature of attitudes that separates them from non-attitudes is generally thought to be that they bear a particular kind of relationship to our reasons.[90] Reasons have been variously characterized as considerations in favor of having or extinguishing an attitude;[91] facts that we respond to with our attitudes and which render them appropriate or inappropriate;[92] and considerations that we bring to bear in answering a certain type of question ('Is p true?'; 'is ϕ choiceworthy?').[93]

For each of the above characterizations, reasons are to be understood to be normative.[94] Facts or considerations regarding physical, circumstantial, or psychological circumstances of a person that motivate or explain her mental states are not reasons in the normative sense unless they also explain why she should have them. Hence, reasons govern which attitudes a person ought to cultivate, maintain, or

---

[90] Or, alternatively, our rational capacity.
[91] Scanlon, *What We Owe to Each Other*, 21; Allan Millar, *Understanding People: Normativity and Rationalizing Explanation* (Oxford: Oxford University Press, 2004) 11.
[92] John Skorupski, *The Domain of Reasons* (Oxford: Oxford University Press, 2011) 37; Joseph Raz, *From Normativity to Responsibility* (Oxford: Oxford University Press, 2011) 5.
[93] Pamela Hieronymi, "The Wrong Kind of Reason," *Journal of Philosophy* 102.9 (2005): 438.
[94] On a number of accounts, reasons are taken to be the most fundamental or basic normative concept. See Raz, *From Normativity to Responsibility*, 6; Scanlon, *What We Owe to Each Other*, 1; Skorupski, *The Domain of Reasons*, 2; and John Broome, *Rationality through Reasoning* (Oxford: John Wiley and Sons, 2013) 36-37. Cf. Niko Kolodny, "Why Be Rational," *Mind* 114.445 (2005): 511-513.

extinguish.[95] Because reasons are normative, we can be asked to account for our reasons for an attitude and we may be open to rational criticism for our attitudes if the considerations or facts that we can produce in support of them are false, or if they don't conclusively favor our attitudes.[96]

The relationship between attitudes and the reasons which favor or oppose them is generally agreed to hold with beliefs, judgments, desires, and intentions. It is also generally agreed that no such relationship holds between reasons and certain other states, such as and sensations, bodily states and processes, brute reflexes, or perceptions, because they aren't the sorts of things that arise or are extinguished for normative reasons. We cannot reasonably be asked to give reasons that recommend them or expect that they should change in response to countervailing reasons. All we can offer for such states are explanations.[97]

The same authors who distinguish attitudes and non-attitudes in this manner generally claim that *emotions* are attitudes.[98] We do tend to think of emotions as the sort of things that we can cite normative reasons for. We can cite reasons that support rather than simply explain them. Imagine that you know a novice teacher who is angry at her

---

[95] Millar, *Understanding People*, 10-15, 42-43; Scorupski, *The Domain of Reasons*, 77; Broome, *Rationality Through Reasoning*, 47; Raz, *From Normativity to Responsibility*, 13-14; Scanlon, *What We Owe to Each Other*, 18; Broome, *Rationality through Reasoning*, 47.

[96] Scanlon, *What We Owe to Each Other*, 21; Millar, *Understanding People*, 39: Hieronymi, "Controlling Attitudes," 49-50; and Pamela Hieronymi, "Reasons for Believing," *Synthese* 161.3 (2008): 371;

[97] See for instance, Raz, *From Normativity to Responsibility*, 85; Scanlon, *What We Owe to Each Other*, 20.

[98] Skorupski, *The Domain of Reasons*, 1-2; Scanlon, *What We Owe to Each Other*, 20; Hieronymi, "Reasons for Believing," 367; Raz, *From Normativity to Responsibility*, 1-5, 48; Millar, *Understanding People*, 14; Alan Gibbard, *Wise Choices, Apt Feelings: A Theory of Normative Judgment* (Cambridge: Harvard University Press, 1990) 36-40.

freshman college students for performing poorly in her class, but you suspect that her anger is unjustified. She presents a case in favor of her anger—she points to facts that support regarding her students as lazy, disrespectful, and entitled. If you're not convinced that the considerations that she presents are accurate and sufficient to justify her anger, you might ask her to consider some additional facts—facts regarding what behaviors she ought reasonably to expect from students at a particular stage of their academic careers (which novice teachers sometimes gauge incorrectly), and whether she might bear some portion of the responsibility for her students' performance for having designed a course on the basis of unreasonable expectations. Perhaps her students are behaving more reasonably and fairly than she thought. If the two of you come to agree that there are decisive reasons for her to abandon her anger toward her students, you expect her to do so, and you would consider her irrational if she didn't. That is, unless circumstances play out in a manner that brings additional reasons to support viewing her students as genuine antagonists to light.

Nonetheless, the idea that emotions could be anything like full blown attitudes has been roundly criticized. A number of philosophers of emotion have recently argued that emotions are in better company with states such as thoughts, perceptions, and other non-attitudes whose contents do not entail reasons or rational commitments.[99] While it's true that our emotions sometimes are altered or extinguished when countervailing facts

[99] Patricia S. Greenspan, "Emotions as Evaluations," *Pacific Philosophical Quarterly* 62 (1981): 158-169; Greenspan, *Emotions and Reasons*, 18; Brady, "The Irrationality of Recalcitrant Emotions," 413-430; and Tappolet, "Emotions, Perceptions, and Emotional Illusions," 1-51.

or considerations come to light, it's also true that our emotions have characteristics that are apparently in tension with attitude-status.

In this chapter I discuss two such features of emotions: their automaticity and their apparent tendency to be recalcitrant. I argue that neither of these features is inconsistent with emotions being attitudes. It is true that many, if not most, of our attitudes arise automatically and non-deliberatively. In response to this worry, I argue that the relationship between attitudes and reasons indicated in the above accounts does not depend entirely or primarily on how they originate. It depends in large part on how they are monitored and managed according to our reasons, and how we assign criticism or blame for them to ourselves and others. Automaticity does not therefore rule out attitude status.

In response to the worry that emotions are especially recalcitrant, or resistant to re-evaluation and conscientious management on the basis of reasoning, I argue three things: First, any kind of attitude can admit of recalcitrant cases. The possibility of recalcitrance is not, by itself, a reason to exclude emotions from the category of attitudes. Second, many apparent cases of recalcitrant emotions aren't *true* cases of recalcitrance. Some emotions may have the appearance of recalcitrance because we judge, for extrinsic reasons, that it would be better for us not to have them. Extrinsic reasons are reasons that favor having or extinguishing a particular attitude *state* regardless of the truth or choiceworthiness of the attitude's *content*. Extrinsic reasons aren't the sort of reasons that are usually considered to have the normative significance that reasons bearing on a state's content do. One's failure to respond to such reasons does not make one's attitude recalcitrant. Also, some emotions may have the

appearance of recalcitrance because they fail to respond to a particular judgment and the reasons that support it, though they are not in fact fixed or unmovable. They can be influenced by reflection. Third, citing evidence on the psychology of emotion regulation, I argue that emotions do ordinarily respond to reflection on our reasons.

1. Emotions and automaticity

One might think that attitudes have the sort of relationship with reasons that they do because they are produced by a rational process. We reason about what to believe, judge, desire, or intend, and the fruits of these activities are our attitudes. Our emotions, too, can be responses to a process of conscious attention to facts or considerations pertaining to our goals and values such as successes, losses, and dangers. However, many of our emotions arise rapidly and automatically, very often without conscious effort of any kind. Thus, you might contend that emotions don't generally originate in the manner that attitudes do. If this claim is correct, then it doesn't obviously make sense to ask a person to justify his or her emotions with reasons expecting a complete and satisfying answer in most cases. After all, if an emotion arises without conscious deliberation about reasons, it can be difficult to articulate the considerations that brought it about. We might also conclude, by this reasoning, that it wouldn't be appropriate to criticize an emotion that arose spontaneously, without intent or identifiable self-direction.

Yet to conclude that emotions are not attitudes because they arise unconsciously and automatically would be a mistake. Many—perhaps most—of our attitudes are formed automatically and non-deliberatively. In fact, our reliance on rapid and

automatic information processing is so extensive that psychologists in the last few decades have come to think of humans as "cognitive misers" whose default mode of operation is to conserve mental resources and to avoid costly correction efforts whenever possible.[100] We reason deliberatively only if we have reason to think that it's necessary, perhaps in an encounter with a truly novel situation or stimulus, or in a situation in which there are reasons to think that our automatic attitude-forming faculties aren't likely to be accurate enough to achieve our goals. Even then, we only engage in reasoning if there are also no other demanding tasks to distract us.[101]

In a manner of speaking, our automatic and non-inferential attitude forming faculties *truncate* the reasons that could ordinarily suffice to justify (or at least rationalize) our attitudes. The reasons they accord with are embedded in an automated procedure that happens unconsciously, but nevertheless brings our background attitudes, our history, and our mental habits to bear on our present situations. If one suspects that such processes are not reliable under one's present circumstances, or when one detects attitude conflicts that demand resolution, one may consciously intervene and attempt to override the outputs of an automatic attitude forming faculty.

What defines our attitudes is not that they are arrived at by a conscious deliberative reasoning process. States that arise without reference to explicitly rehearsed

---

[100] See Keith E. Stanovich, "Distinguishing the Reflective, Algorithmic, and Autonomous Minds: Is it Time for a Tri-Process Theory?" in *In Two Minds: Dual Processes and Beyond*, edited by Jonathan St. B. T. Evans and Keith Frankish (Oxford: Oxford University Press, 2009) 69. Stanovich describes the "cognitive miserliness" trend in thinking about human information processes before disambiguating several features of our miserliness.

[101] See Conor McHugh, "Epistemic Responsibility and Doxastic Agency," 132-157 for a helpful discussion of self-monitoring and its implications for attitudes produced automatically, without reasoning or reflection.

reasons nonetheless bear the right sort of relationship to our reasons to count as attitudes because we—the proprietors and managers of them—can cultivate, monitor, and modify them in a way that respects our reasons to a greater or lesser extent. Our failures to respect our reasons are rational failures.

Although there are difficulties associated with articulating whatever reasons one responded to in forming an attitude on the basis of an automatic or non-inferential process, it us not unreasonable for us to ask one another to justify such attitudes. Consider a fairly simple perceptual belief. If there is a yellow curtain before me, I'll tend to believe that there is one though I typically call no reasons to mind explicitly in so-believing. There is usually no need. However, if I'm asked to give an account of my reasons for believing as I do—say, as an eye witness to events that have come into question by an interlocutor who doubts that my belief is correct—I will likely be able to present facts or considerations in favor of it. Our perceptual beliefs are the products of processes that are largely opaque to us, but we can support them with facts and considerations such as the reliability of the processes under the present conditions. Perhaps I'll cite the quality of lighting in the room, the distance between myself and the objects in question, and my ability to correctly identify curtains and fabric colors and configurations under these conditions. We can also supply our understanding of the concepts applied in the perception. I can demonstrate my understanding of the relevant concepts by citing facts about the differences between curtains and other household linens, and so forth. In fact, none of these reasons were immediate, overt causes of my belief. I didn't explicitly consider them in forming it. Still, they are the sorts of facts that support my attitude and might justify my maintaining it.

It may be the case that, when asked for our reasons, we wind up at a loss because we have neither knowledge of the process, nor any particular reasons to think that it's reliable under the present conditions. Or, we might produce reasons that have little to do with the actual process and its reliability. For example, in Crandal and Calderwood's study of neonatal nurses, experienced nurses demonstrated finely tuned intuitions about how the symptoms of sepsis present themselves in infants. The nurses in the study were able to quickly and accurately identify patients with the condition and implement proper treatment and care using intuitive judgments rather than slow, cognitively costly deliberation about the reasons present in a particular case. These intuitions were potentially lifesaving, as early diagnosis is important for successful treatment of the condition. The nurses were able to correctly diagnose it before the infection that was causing it had advanced to the point at which laboratory tests could detect it. Although the nurses had stable, reliable perceptual-recognitional dispositions supporting accurate diagnoses, their intuitions weren't normally accompanied by well-integrated explicit knowledge of the sorts of considerations that made their intuitive judgments successful. When asked to identify the markers of sepsis that supported their judgments regarding which infants were at high risk, they weren't able to do so accurately or sufficiently. Crandall and Calderwood speculated that this might have had something to do with the fact that medical journals and manuals listed symptoms of sepsis in adults, though the condition apparently presented differently in infants. The nurses' practical knowledge outstripped their formal training, and so the reasons that

were readily accessible to them in justifying their diagnoses due to their training weren't actually relevant to their patient's cases.[102]

Crandall and Calderwood's study can illustrate two important points about the relationship between attitudes and reasons that stem from our reliance on automatic and non-deliberative processes. First, there may be attitudes for which we have no articulable or explicit reasons. If a nurse had difficulty articulating *any reasons at all* for his judgments about which infants have sepsis, it would still be wrong to say that he has no attitudes about sepsis in infants. In fact, he does regularly form judgments about sepsis in infants in his care. The perceptual-recognitional skills that he deploys in forming these attitudes are, nevertheless, opaque to him and he doesn't have a theoretical grasp of his skill with which to justify his diagnoses when they come into question. It's possible that he could produce explicit reasons for these attitudes even if he is not now in a good position to do so.[103] Second, there may be attitudes that we hold for reasons that diverge in important ways from the ones that we are actually able to articulate. In the study, most nurses could present reasons to support their diagnoses, even though they might only be able to offer incomplete or inaccurate accounts of the sorts of reasons that they actually brought to bear in forming the relevant intuitive

---

[102] B. Crandall and R. Calderwood, "Clinical Assessment Skills of Experienced Neonatal Intensive Care Nurses (Final Report)," (Yellow Springs, OH: Klein Associates Inc., 1989), Prepared under contract 1 R43 NR0191101 for the National Center for Nursing, NIH.; Op cit. Gary Klein, "Developing Expertise in Decision Making," *Thinking and Reasoning 3.4* (1997): 344-5

[103] Note the difference between a refusal of a request for reasons and an answer such as "no reason." The answer "no reason" is not a refusal of the request for reasons. On the contrary, it implies that the question "what reasons do you have" does, in fact, have application. It implies that one could have reasons, although in this case one doesn't. See Hieronymi, "Reasons for Believing" 360; and G. E. M. Anscombe, *Intention* (Ithica, N.Y.: Cornell University Press, 1976) 57.

judgments. Their reports of their reasons, in such cases, are open to criticism. It can sensibly be pointed out that something presented in the guise of a reason isn't factual, or isn't relevant to a particular case. I conclude that automaticity is often correlated with lack of self-knowledge regarding our reasons, but this need not rule out something's being an attitude.

Like any of our attitudes, our emotions can either arise as responses to conscious reflection on our reasons, or they can arise automatically and non-deliberatively. No matter how they originate, or how explicit a grasp we have of the reasons that our emotions track, our emotions are the sorts of states that can be had for reasons. Just as we are with our automatic beliefs, judgments, desires, and intentions, we are the proprietors and managers of our emotions, and we can manage them in a way that respects our reasons to a greater or lesser extent.

Even when our emotion faculties operate automatically, they embody our past experiences, attitudes, and mental habits. We develop emotional dispositions toward particular kinds of stimuli by developing a structure of implicit memory and automated procedure that bypasses slower reasoning and reflection processes. This allows us to react to value-relevant features of our environment in a manner that is both efficient and sensitive to our reasons. When we suspect that an emotion that formed rapidly and automatically in response to a situation is inappropriate given the complete set of relevant facts or considerations to which we have access, we can consciously intervene to change a present emotional response. In turn, conscious interventions shape our

dispositions so that our automatic emotion faculties will display greater accuracy in the future.[104]

In the next section, I address the issue of emotional recalcitrance. Our emotions do not always respond to our judgment-oriented reasoning efforts, at least not obviously or immediately, if at all. One might worry that this suggests that our emotions are not, like our (other) automatically-produced attitudes, the sort of thing that we are genuinely able to intervene upon or make adjustments to. In sections 3 and 4, I develop a case for the view that emotions do respond to reflective interventions, even though they don't always respond to the outputs of our judgment-oriented reasoning processes. Our reflective capacities consist in a larger tool-kit than reasoning alone. There are other reflective capacities that we can use to greater effect in the management of our emotions.

2. Recalcitrant emotions

The claim that emotions are related to our reasons in the manner of an attitude is in apparent tension with the fact that emotions seem to be highly resistant to change in light of our judgments about our reasons in some cases. For something to be an attitude, it needs to be responsive to our reflective capacity because states that aren't responsive to reflection cannot be controlled or managed in accordance with our reasons. Judgment is an important component of reflection. In normative philosophy, it

[104] Anett Gyurak, James J. Gross, and Amit Etkin, "Explicit and Implicit Emotion Regulation: a Dual Process Framework," *Cognition and Emotion* 25.3 (2011): 405-407.

is not uncommon to assume, at least implicitly, that judgment-oriented reasoning is *the* reflective capacity.

I begin my discussion of this issue with a careful examination of the phenomenon of emotional recalcitrance. I start by pointing out that the possibility of recalcitrance does not rule out something's being an attitude. Nevertheless, I will argue that many apparent cases of emotional recalcitrance are not cases in which the emotion actually fails to respond to normative reasons. Some apparent reasons fall out of the purview of normative reasons because they aren't appropriately related to the attitude's truth or choiceworthiness. Failure to respond to such reasons does not make an emotion truly recalcitrant, though it can give an emotion the false appearance of recalcitrance. Other apparently recalcitrant attitudes simply fail to respond to the particular reasons that moved us to judge differently. They nevertheless would respond to other reflective strategies that bring the reasons the emotion is (or would) respond to into focus, for reasons discussed in section 4. Judgment-oriented reasoning is by no means our only reflective capacity. Emotions that fail to respond to judgment-oriented reasoning, but would respond to other varieties of reflection are not truly recalcitrant.

i. True recalcitrance

Emotions do sometimes genuinely fail to respond to reasons, exhibiting true recalcitrance. We might have a fear, for example, that is genuinely unsupported by any sort of non-negligible danger considerations, but that will not be moved by any facts whatsoever regarding the lack of danger. Even so, recalcitrance is a possibility for any attitude-type. Here is an example of a recalcitrant belief: Gretchen believes that her

mother is in good health. Then, while driving through town one day, she sees her mother leaving a cancer treatment center. Later, she finds drugs in the drawer of her mother's nightstand that are used to treat cancer patients. Her mother's phone rings and the caller ID indicates that the call is coming from a doctor's office. Gretchen recalls that her mother's mood and energy level have recently declined in a dramatic fashion. Although Gretchen faces mounting evidence that her mother is no longer in good health, she continues to believe that she is. Recalcitrant attitudes, like Gretchen's, might result from motivated believing. Gretchen doesn't want to accept her mother's mortality because she depends on her mother's love, support, and guidance. She thus brings herself to believe in a more optimistic explanation of her evidence rather than a much more likely one, even as her evidence for the—to her—unthinkable conclusion continues to accumulate. Her belief may become so entrenched that any considerations to the contrary make no difference to her attitude.

Cases like this one need not reflect on the reason-responding nature of our attitudes generally speaking. Our beliefs, desires, judgments, and intentions can be recalcitrant, but they are typically moved by reasons. Our emotions are too. The simple fact that attitudes of a type are sometimes recalcitrant is no reason to think that they are not typically moved by all (or at least many) of the relevant facts or considerations.

Still, emotions might seem to be especially recalcitrant because they are unresponsive to our reasoned judgments on a fairly regular basis. It's not unusual for one to experience episodes of shame while at the same time one judges that one ought to be proud of oneself. It's likewise fairly commonplace for one to experience jealousy though one judges that there is no significant threat to one's valued relationships. If

emotions are recalcitrant more than beliefs, judgments, and intentions are, one might think that they simply aren't appropriately connected to our reasons in the manner of an attitude. Our attitudes, it's natural to assume, are guided and corrected by judgments about relevant facts and considerations.

I respond to this worry with a discussion of the following case: Francis is extremely unhappy with his career as a college professor. The work has begun to feel toilsome and dull and the love for his discipline that used to keep him going is starting to slip away. Nevertheless, he judges after careful consideration that he ought to continue in his profession and that his situation thoroughly warrants happiness. He is doing well by professional standards, he believes that his work is contributing to social goods, and he is reasonably well-compensated for his work. These are just the things that he takes to be of the highest value to him in his work life. Moreover, he reasons, a career change could be costly. As happiness is an emotion that implies general wellbeing and value-satisfaction, Francis may be well-justified in a judgment that his unhappiness in inappropriate to his circumstances. For the present, he stays in his career, but in spite of what he takes his reasons to be, he cannot bring himself to be happy in his work.

Francis is experiencing a persistent emotion/judgment conflict. It is not, however, necessarily correct to conclude on this basis that he is experiencing emotional recalcitrance. There are two ways that an emotion might appear to be recalcitrant due to a conflict with judgment even though the emotion is responsive to reasons. First, at least some cases of apparently recalcitrant emotions are cases in which the reasons for or

against the emotion are what Pamela Hieronymi calls "extrinsic" reasons.[105] They aren't reasons pertaining to the attitude's content. They are pragmatic reasons in favor of forming or extinguishing the attitude. The rationality of our attitudes does not depend on our responding to such reasons.

Other emotion/judgment conflicts involve genuinely conflicting reasons that do pertain to the attitude's content. However, an emotion that conflicts with a judgment need not be immovable. The emotion may respond to forms of reflection other than judgment-oriented reasoning. Francis's case can illustrate both of kinds of apparent, but false recalcitrance, which I discuss in the sections that follow.

ii. Attitude conflicts and the wrong kinds of reasons

Some apparent cases of emotional recalcitrance are due to the fact that the considerations for or against adopting an attitude are the wrong kinds of reasons: considerations that that justify the *adoption* of the attitude state that aren't, at the same time, considerations that pertain to the appropriateness of the attitude's content. Our attitudes are not typically responsive to these reasons, nor should they be.

In her work on attitudes and the will, Pamela Hieronymi introduces a helpful distinction between types of reasons for an attitude: intrinsic reasons and extrinsic reasons. Hieronymi introduces the distinction as a solution to an apparent problem with the simple analysis of reasons as considerations or facts bearing on whether to have an attitude. Such an analysis fail to distinguish the right kind of reasons—reasons that are normative in the sense that is relevant to our attitudes—from the wrong kind of reasons.

[105] Hieronymi, "The Wrong Kinds of Reasons," 452.

The right kinds of reasons are *constitutive* reasons, which bear on an attitude's content. Since belief in *p* is an attitude of assent toward *p*, reasons in favor of *p*'s being true are *constitutive* reasons for believing *p*. We should expect our beliefs to be responsive to such reasons. The wrong kinds of reasons are *extrinsic* reasons, which bear on whether it would be good to be in a particular attitude state, regardless of its content. If something is a consideration in favor of believing *p* because believing *p* would lead to a practically advantageous outcome, it's an extrinsic reason.[106]

We shouldn't expect our attitudes toward p to turn on considerations regarding the advantages that having such attitudes would proffer. Such reasons aren't what normally would (or should) move us to adopt an attitude. For example, genuinely believing that I am the best philosopher of emotions might have certain advantages for me. I'd be more likely to submit my work to good conferences and journals, and I'd appear more confident toward my work, inspiring greater confidence from others. These are reasons in favor of adopting the attitude, but they are not reasons for the attitude itself. It's doubtful that I could bring about such an attitude just by reflecting on the fact that it would be good for me to have it. I would likely have to seek constitutive reasons to believe it in order to bring myself to believe it, which might require me to engage in some self-deception and wishful thinking. Accordingly, the extrinsic reasons that support my attitude in such a case are not straightforwardly rationally or epistemically normative in the same sense that my constitutive reasons are, since my extrinsic reasons recommend the attitude regardless of whether there are constitutive reasons against it.

[106] Ibid., 452.

They don't obviously justify our attitudes in a manner that shields us from criticism on the basis of contravening constitutive reasons.

What Hieronymi's distinction between constitutive and extrinsic reasons highlights is the fact that we should not expect our attitudes to be responsive to just any kind of consideration that bears on an attitude. We should rather expect our attitudes to be moved by constitutive reasons for or against an attitude. Emotions, as reason-responding intentional attitudes, also admit of the both constitutive and extrinsic reasons.

For example, constitutive reasons for anger, an attitude that implies acknowledgement of a slight, are reasons bearing on the question of whether one (or someone one cares about) has been slighted. Practical reasons to be angry—to cause fright to those who might worry that your anger will result in retaliatory actions against them, or to please others who expect your solidarity in their own anger—are extrinsic reasons. We should not expect our emotions to be moved by such reasons, though they might provide us with compelling reasons to try to induce them. As with our other attitudes, extrinsic reasons are not sufficiently germane to the attitude to properly or appropriately move one to have it regardless of whether there are decisive constitutive reasons not to.

Returning to Francis's case, some of the reasons that form the basis for his judgment that he ought to change his emotion are extrinsic reasons. Constitutive reasons to be happy or unhappy with one's situation must pertain to the question of whether the situation is, itself, good or valuable to one. Wanting to avoid the costs associated with moving from a moderately successful career into a new one in which one might find

less success is a reason for Francis to try to bring it about that he is happy with his career, but it's not a reason for him to be happy. Thus, to fail to be moved by such reasons does not make his unhappiness recalcitrant.

To bring it about that he is happy, Francis can seek constitutive reasons to be happy. This is the tack he pursues when he focuses on the good-making features of his work: his professional success and his adequate compensation. However, though he judges these reasons sufficient to warrant his happiness, he may not experience happiness on the basis of these reasons. In the next section, I will argue that this need not imply that his emotion is not responding to reasons. His judgment may have been formed on the basis of a narrower or differently weighted set of reasons than his emotion.

iii. Attitude conflicts and opposing reasons

Human beings are not omniscient with respect to their own reasons. Accordingly, we do not access every consideration or every relevant attitude in our entire attitude set every time we make a judgment. The potentially important implications of this fact have been noted in philosophical debates about akrasia. Akrasia—a failure to do what one judges best— often involves emotion/judgment conflicts. Nomy Arpaly's Susan case is suggestive:

> If Susan, being a human being, deliberates imperfectly and forgets to consider a thing or two before concluding ''all things considered, I should marry Todd,'' it is still true that *her best judgment* tells her that she should marry Todd.[107]

[107] Nomy Arpaly, "On Acting Rationally Against One's Best Judgment," *Ethics* 110.3 (2000): 490.

We can imagine Susan persisting in her judgment, but *being unhappy* as she prepares for the wedding. She anticipates *feelings of regret* for following through with it, and she is *hoping* for a convenient way out of the situation. Eventually, she acts on her feelings and calls off the wedding. Her action is akratic because she failed to act on what she judged best.

Her judgment that she ought to marry Todd, we're told, is based on a too narrow set of considerations, though they may seem comprehensive to her when she forms her judgment. Perhaps she is overvaluing her current feelings of love and loyalty and her expectations for financial stability in the marriage. Yet we can imagine that part of her (her emotional sense) recognizes that she implicitly values things that she does not expect to find in the relationship in the long term—things such as respect and equal partnership. Such considerations may not come to mind when Susan judges that she should marry Todd, but they are salient when it comes to such important life decisions. Given her entire set of attitudes, values, and commitments, she may have overriding reasons not to marry Todd. Therefore, although Susan acted against her judgment when she called off her wedding, she may very well have acted in accord with her reasons.

Cases like these remind us that acting on the basis of our best reasons might not be the same thing as acting on the basis of our judgments.[108] Our reasons are determined by the whole set of attitudes and values that we have. We can imagine that Susan had a total set of commitments that recommended against what she judged she ought to do on the basis of careful consideration. Her judgment may have been produced by a reasoning process in which she neglected some important considerations and

[108] Ibid.

inappropriately weighted others. A more careful consideration of her complete set of commitments might have yielded the judgment that she should not marry Todd after all.

Ordinary human agents do not always perform a careful reasoning processes to form attitudes. Even when one does reason carefully, one's judgment is subject to certain flaws and limitation. Assume that Susan's error was innocent. Her judgment was not due to wishful thinking, recklessness, or hurried and sloppy calculations, even though it was, ultimately, ill-conceived. The human judgment faculty operates with limited information due to the low capacity of conscious thought.[109] Thus, when we engage in deliberative judgment, we're liable to commit inadvertent errors of omission. We focus on the reasons that seem most salient to us at the time, but we fail to include other relevant reasons that might have tipped the balance in favor of a different judgment.[110] We are also disposed to unknowingly employ suboptimal weighting schemes when we deliberate, especially if the subject matter is unfamiliar or simply outside of our expertise. The values that we think are important to us when we make a judgment under such conditions may not be as important to us as we think. Other values may actually determine how satisfied we are with our choices in the long run.

These potential pitfalls of deliberation have been illustrated in studies about reasoning and post-choice satisfaction. For example, Wilson et al. conducted a study in which college students were asked to rate their liking for a number of posters: A print of Monet's *Nymphé*as, a print of Van Gough's *Les Iris Saint Rémy*, a cartoon depiction of

---

[109] Ap Dijksterhuis and Zeger van Olden, "On the Benefits of Thinking Unconsciously: Unconscious Thought Can Increase Post-Choice Satisfaction," *Journal of Experimental Social Psychology* 42 (2006): 627-8.
[110] Timothy D. Wilson et al., "Introspecting about Reasons can Reduce Post-Choice Satisfaction," *Personality and Social Psychology Bulletin* 19.3 (1999): 332.

animals in a hot air balloon, a picture of a cat balancing on a rope captioned, “Gimme a Break,” and a picture of a cat standing by a fence captioned “One Step at a Time.”[111] Some students were tested in a “reasons condition” group. These students were given a questionnaire encouraging them to write down reasons for liking or disliking each poster. Then they were asked to rate the posters. Other students were tested in a “no reasons” group. These students were given a questionnaire about topics unrelated to the posters such as their majors and their career plans before they were asked to rate the posters. After rating the posters, the students in both groups were given their choice of poster to take home.

The control subjects, who were not asked to enumerate reasons for liking or disliking the posters, tended to rank the art posters a great deal higher than the humorous posters. The test subjects, who listed reasons for their preferences, rated the art posters lower than the “no reasons” subjects and the humorous posters higher than the “no reasons” subjects. Nearly all of the control subjects took an art poster home. “Reasons condition” test subjects were much more likely to take a humorous poster home than their non-reasoning counterparts. After several weeks had passed, experimenters contacted the test subjects to ask questions designed to determine how satisfied they were with their choice of poster: Whether they kept it, whether they hung it up, how they would rate their liking of it on a 10 point scale, and how much they’d be willing to sell it for. Subjects who reasoned about the likeable qualities of the posters

[111] Ibid., 333.

were less satisfied with their choices, especially the ones who chose the humorous posters.[112]

As a result of the study, Wilson et al suggest that in at least some circumstances, we form superior attitudes and make better decisions when we refrain from explicit reasoning and 'go with our guts.' Judgment has limitations when it comes to accessing our entire set of reasons and weighting them appropriately, so it's not always good for us to follow our judgments. We'll sometimes end up with superior decisions that better accord with our entire set of subjective reasons when we act *against* our best judgment.

Susan's decision to marry is one that will have bearing on almost every facet of her life. The rationality of the decision involves many considerations and a very complex weighting of values. Susan would do better to go with her guts given that her emotional response is in fact the correct one in the situation. However, Susan, who has judged that she ought to marry Todd, has failed to consider the values, goals, and beliefs that she has that would support abandoning her plan to marry Todd. It would be unwise for her to simply go with her guts. Our emotional faculties can, after all, produce ill-supported attitudes, just as our judgment faculties can. When we're reacting emotionally to our expectations about future events, we can fixate on an unduly narrow set of considerations or weigh some considerations inappropriately.[113] As far as Susan can tell, her emotions might be responding to a narrow or inappropriately weighted set of fear- and uncertainty-considerations that are truly outweighed in a grander scheme of

---

[112] Ibid., 334-5.
[113] George Loewenstein, "Affective Regulation and Affective Forecasting," in *Handbook of Emotion Regulation*, edited by James J. Gross (New York: Guilford Press, 2007) 187.

happiness-considerations that make marrying Todd the best thing for her to do. Nevertheless, the fact that Susan is feeling inclined not to do what she's judged she ought is a clue that she might need to reconsider her judgment that she ought to marry Todd. We can make better decisions if we take precautions to ensure that we carefully consider our reasons when we're experiencing inner conflicts.[114]

As I argue in section 4, truly careful consideration of one's reasons in an emotion/judgment conflict is not a simple re-hashing of one's original deliberative reasons. It requires employing other reflective strategies such as imagination and perspective taking to probe one's values, the reasons that they generate, and the relative weighting that one's reasons ought to be given.

Susan's case is very like Francis's. Francis's unhappy emotions may very well put him in akratic mood, and his akratic mood may result in actions that go against his judgment: quitting his job and looking for a new career that will make him happy. If he acts in accordance with his judgment and stays, he may eventually find himself happy and fulfilled in his job again. This depends, in large part, on how well his judgment took into consideration all of the relevant factors concerning what, to him, makes something worth being happy about. The reasons he judges sufficient for being happy, stated in the case, are that he's professionally successful, his work is valuable, and he is reasonably compensated. Those sorts of things tend to be important to our happiness in our professional lives, and Francis may be no exception. However, there are conceivable

[114] See Lisa Bortolotti, "The Epistemic Benefits of Reason Giving," *Theory and Psychology* 19.5 (2009): 624-645. Bortolotti provides arguments against the claim that it would be a good policy to "go with your guts" instead of engaging in reasoning and reflection, even though our attempts at reasoning and reflection can sometimes lead us astray.

reasons why those things would not be enough to make one happy in one's career. For example, one might value collegiality and community-feeling with one's coworkers very highly. In a workplace that is cold or hostile, one might be very unhappy with one's situation in spite of its other positive features. Though one might maintain, in judgment, the notion that workplace social-climate isn't as important as professional success for happiness in one's career, one may still feel unhappy, suggesting that emotionally, one ranks such considerations more highly than one does in judgment.

In a case such as Francis's then, it's not necessarily that his emotion isn't responding to reasons. His emotion may simply not be responding to the reasons that he's based his judgment upon. Emotions and judgments may operate on the same plane of reasons, but they may not yield consistent attitudes due to the way that we identify, organize, and assess reasons in the formation of these attitudes. We should expect to see attitude discrepancies between emotions and judgments on some occasions.

iv. Emotional recalcitrance reconsidered

Emotions are often thought to be too recalcitrant be attitudes because they appear to be so often unaffected by our judgments about our reasons. In section 2 parts *i.* through *iii.*, I explained what it means for an attitude to be recalcitrant in the relevant sense by distinguishing true recalcitrance from merely apparent recalcitrance. In brief, an attitude is recalcitrant if it doesn't respond to reasons in favor of or against the attitude's content. An attitude need not be truly recalcitrant simply because it doesn't respond to extrinsic reasons—reasons that make having the attitude worthwhile, but don't have any bearing on whether the attitude is, itself, reasonable or rational. Also, an attitude need not be truly recalcitrant simply because it doesn't match a judgment about

what our reasons support. Emotions, I'm suggesting, are not as prone to being recalcitrant as they are sometimes claimed to be.

If an emotion doesn't respond to judgment-oriented reasoning, the question remains: How is one to exercise reflective agency over it? If something is an attitude, we can ordinarily expect it to be affected by the activity of reflection. If an emotion doesn't respond to judgment-oriented reasoning, are we left with no recourse—no way to resolve our emotion/judgment conflicts? The answer to this question is no. As Valerie Tiberius reminds us in *In Defense of Reflection*, judgment-oriented reasoning is part of a larger repertoire of tools and capabilities that we utilize when we reflect. She writes,

> …analyzing reasons—the kind of reflection that is studied in much of the psychological research—is not obviously the most important kind of reflection. Indeed, the value of analyzing reasons (if it has value) seems to depend on its role in a broader form of judgment that considers the situation holistically and with attention to context. Two reflective methods that come to mind when we think about this form of reflection are imagination and perspective taking, whereby we can try to picture what things will be like for us if we choose one way or another or put ourselves in the shoes of other people to see what things are like from their perspective.[115]

Tiberius maintains that although judgment-oriented reasoning is an important tool for reflection on our reasons, goals, and values, there is also a broader range of reflective capabilities that we can—and should—utilize when we're assessing our reasons. In noting that there is more to reflection than judgment-oriented reasoning, we open up the possibility that other reflective strategies might succeed to effect change in our emotions where judgment-oriented reasoning has failed. I turn to this possibility in

[115] Valerie Tiberius, "In Defense of Reflection," *Philosophical Issues* 23 (2013): 233.

section 3. I discuss the effects of a variety of reflection strategies on our emotions using Francis's case to illustrate.

3. Adjudicating emotion/judgment conflicts

I've suggested that Francis's emotions and judgments about his career situation may both be reason-responding even though they conflict. They conflict either because they are responding to different reasons or because they are responding to the same reasons under different weighting schemes. If his emotions and his judgments are reason-responding, we expect them to be affected by the activity of reflection. Here, I survey some reflective tools that Francis could utilize to adjudicate his emotion judgment conflict.

Before I begin this survey, it bears repeating that although it's commonly assumed that when a conflict of attitudes arises, we ought to be guided by judgment, this isn't necessarily the best strategy for reconciling our attitudes to accord with our reasons. Judgment, I have argued, may operate on suboptimal value weighting schemes. It may also exhibit too narrow a scope. Even when we are exercising care in our deliberations, we tend to focus on easily accessible considerations, leaving out other salient, but less obvious considerations. Emotions, too, may exhibit suboptimal weighting schemes and limited scope. Our emotional sensibilities can be poorly calibrated so that we respond to value-relevant objects and states of affairs in a manner that is over- or under-sensitive to certain kinds of considerations. We may then experience emotions toward objects and situations that we genuinely shouldn't.

How, then, should Francis resolve his emotion/judgment conflict? To adequately adjudicate his emotion-judgment conflict, it's not sufficient for him to simply engage in further judgment-oriented deliberation. Such an exercise would not necessarily correct for the faults of his judgment. Nor would it afford an adequate measure of regard for his emotions. Adjudicating emotion/judgment conflicts requires employing reflective strategies that can help us probe our values and recalibrate our weighting schemes. It requires that we use reflective strategies that are capable of expanding the purview of both emotion and judgment so that all of the relevant reasons can be brought to bear in the revision of either or both attitudes.

Reflective activities that involve careful inspection of an emotional object or situation can, when wielded effectively, have the power to interrupt emotions that do not suit their objects and draw out emotions that respond to the right sorts of reasons. The same reflective activities can draw attention to the fact that there may be reasons one has ignored or inappropriately weighted in judgment. In the paragraphs to follow, I elaborate on a sampling of reflective strategies and their facility in changing our attitudes, including our emotions, to better accord with our reasons. I also discuss several non-reflective strategies for attitude management, and particularly the non-reflective strategies studied in research on emotion regulation. I argue that these strategies ought to be wielded with caution. When reasons-reflection is possible, it is both more effective and less likely to yield false or otherwise inapt attitudes.

i. Attention redirection

One reflective strategy that might be useful in adjudicating an emotion/judgment conflict is attention redirection. Attention redirection draws out new facts or considerations that may support either the emotion or the judgment. It is also a demonstrably effective strategy for emotional change. If an emotion is truly inappropriate, attention redirection can interfere with automatic, preconscious attention biases that maintain positive feedback loops between emotions, perceptions, and concept deployment.

Attention biases help maintain emotions by facilitating the detection of emotion relevant stimuli. For example, a person who is experiencing social anxiety identifies frowning faces more quickly than a person who is not anxious. She is also more likely to identify neutral faces as frowning (a hostility cue) than a person who is not anxious. Or, a person who is anxious may avoid attending to faces altogether. Often, the same person will alternate between hypervigilance (seeking faces to scrutinize) at one moment and avoidance (not attending to faces at all) at another. Both contribute to the emotion's continuation by increasing false negatives in hypervigilance and preventing more careful scrutiny of the stimuli in avoidance.[116] Attending to stimuli that are incongruent with an emotional evaluation and the perceptions and conceptions that they promote can interrupt the feedback loop. This creates opportunities to shift one's focus to other features of the object or situation that would elicit more apt emotions.[117] The

[116] Pierre Philippot et al., "Cognitive Regulation of Emotion: Application to Clinical Disorders," in *The Regulation of Emotion*, edited by Pierre Philippot and Robert S. Feldman (Mahwah, New Jersey: Lawrence Erlbaum, 2004) 84-5.
[117] Ibid., 81-3.

same activity creates opportunities to reassess the reasons that were the basis of one's judgments that conflict with the emotion.

In Francis's case, instead of allowing his emotional state to determine which aspects of his career he focuses on (or avoids) and which concepts he employs when thinking about them, he could seek out aspects of his job that don't fit a schema of unhappiness—small pleasures, positive relationships, and so forth—that could possibly break the emotional feedback loop and allow him to reframe his outlook on his situation and produce a change in his emotions. However, if Francis's unhappiness is really warranted, he might not find enough of the relevant sorts of considerations to produce a change in his emotional response. The exercise might provide him with better reasons to change his judgment instead.

ii. Information elaboration

Information elaboration is reflection on the same elements of the object or situation that cause or maintain the emotion. Overgeneralization may facilitate recurring patterns of thought that are congruent with the emotion. In turn, this prevents one from attending to elements of the situation that are incongruent with the emotion. In one study, participants were put in an anxiety inducing situation: they were asked to give an oral presentation for which they would be evaluated. Before the presentation, they were given one of the following tasks. Members of the first group were asked to do a specific thinking task. They answered increasingly specific questions regarding their worries about their upcoming presentations. Members of the second group were asked to do a general thinking task. They answered questions about their general impressions

regarding their situation. Members of the third (control) group were given a distraction task. They were asked to come up with antonyms for several non-emotional words. As a result, members of the first group, who were asked to do specific thinking about their situation exhibited significantly decreased anxiety. Members of the second (general thinking) group exhibited increased anxiety and members of the third (control) group exhibited stable anxiety levels.[118]

Information elaboration can disrupt an emotional response and inhibit feedback loops by promoting clearer thinking about the matter. By attending to the details of the emotional situation or object, one reduces focus on a set of overgeneralized considerations that reinforce the emotion.[119] By the same token, information elaboration can produce information that has bearing on one's judgment that is in conflict with one's emotion.

Francis can meditate on the features and situations at his job that seem to contribute most to his unhappiness. If he directly confronts the things that are making him unhappy, he may be able to draw himself away from the excessively negative thoughts that maintain his emotional attitude and find considerations which give it a less negative aspect. For example, if he is unhappy with the way his classes are going, and he meditates on the things that strike him to be bad about them, he may find that he's exaggerated their badness in his unhappiness. The students haven't been truly as lazy, mean, and unteachable as he's been making them out to be. The situation may then take on a different emotional significance for him. On the other hand, if close inspection of

[118] Ibid., 86-91.
[119] Ibid., 83.

the situation reveals that his classes lately have been truly and dreadfully unsuccessful, he may find reasons to reassess his previous judgment about the worthiness of his career at present.

iii. Imagination

Imagination is an effective tool for resolving emotion/judgment conflicts because in imagination one can construct a representation of future situations and reflect on how these constructs might bear on our goals and values. This creates possibilities for attention redirection and information elaboration. With imagination, we can explore a variety of contingencies with as much detail as our knowledge makes possible, and we can explore them from a variety of perspectives. We can even explore what it might be like to embody another person's perspective in the situation.[120]

If, instead of reflecting on facts in the abstract, Francis vividly imagines leaving his job and embarking on a new career venture, he can put the facts into a richer context that brings values into focus. Then, he can probe his emotional sense: Would he miss his job, or would he feel relieved to be rid of it? Would he be excited to try a new career, or would he feel regretful for leaving his old one? He might discover something

[120] There will of course be gaps in our knowledge regarding what the future situations that we imagine will actually be like, and those unknowns might cause emotional responses in future circumstances that we're not able to anticipate in the present. For instance, researchers asked individuals to predict how happy or unhappy they'd be as a paraplegic. Non-paraplegics tended to predict much greater unhappiness than actual paraplegics tend to report. This might be explained by the accessibility of considerations regarding what we can now do, but wouldn't be able to do as paraplegics. Actual paraplegics have a better grasp of the positive considerations (what they can do in spite of the condition) to offset the negative ones. See Loewenstein, "Affective Regulation and Affective Forecasting," 187.

about his values that he missed before. He may realize that his current job does still hold meaning for him that his emotional sense wasn't previously capturing. He may also gain insight into the specific features of his job that bug him and that he would be especially glad to leave behind. The exercise might result in a change or tempering of his emotion. It might, alternatively, or in addition, change his judgment.

iv. Perspective taking and perspective shifting

Perspective taking is another kind of imaginative exercise that can aid attention redirection and information elaboration. In perspective taking, one displaces the self and imaginatively enters the perspective of another to try to understand his or her evaluative standpoint. One focuses on the features of the situation that person would focus on and one attempts to see the value in the things that they value. Gaining an understanding from another perspective can help one to alter one's own evaluative standpoint. Francis, for instance, might try to imaginatively enter a happy colleagues' perspective. From this new perspective, he might notice some feature of his job that he values, but that he underappreciated when forming either his emotion or judgment.[121]

Perspective shifting, like perspective taking, requires altering one's evaluative standpoint. Unlike perspective taking, a perspective shift requires reorienting one's evaluative standpoint around some value that already has a place—perhaps even an important one—in one's own value-set. In the present situation, however, that particular value has ceded the foreground to other values and goals that are most accessible in the

[121] See Jonathan Haidt, "The Emotional Dog and its Rational Tail: A Social Intuitionist Approach to Moral Judgment," *Psychological Review* 108.4 (2001): 819, 829. Haidt discusses role taking's place in moral reasoning.

present situation. For example, in one's role as parent, one may have a set of goals and values revolving around the welfare of one's children. In one's parenting role, other important self-regarding goals and values may recede from view. At times, those self-regarding goals and values may return to the fore. Each set of goals and values might promote different sets of expectations, judgments, and emotional responses. As a parent, one prioritizes the welfare of one's children above all. As an individual, one has personal goals and values which may include a flourishing family as well as other goals (educational, career, spiritual, travel, etc). Deliberate shifts between these goal and value-organizing and orienting perspectives might promote better goal fulfillment and more accurate and appropriate emotional responses overall.[122]

A perspective shifting strategy could be useful to Francis if there is an important value that he's backgrounded in thinking about his career and its overall importance to him. Perhaps Francis didn't originally go into philosophy because he cared about having an interesting and personally fulfilling career. He went into it because he was pretty good at it and it could afford him a moderately comfortable salary in a cozy college town, respectability in the community where he lives, and a flexible, though busy, schedule that allows him to do other things that he enjoys every now and then. Perhaps foregrounding these goals and values could make the sorts of things that are currently making him unhappy about the job itself seem less important than they seemed when he was thinking from his perspective as a career-holder. This might result

[122] See Valerie Tiberius, "Wisdom and Perspective," *The Journal of Philosophy* 102.4 (2005):173-4. Tiberius suggests that we move between broader and narrower perspectives as we pursue our various goals. Her discussion includes comments on the relationship between practical perspectives and emotions.

in a modification in his emotion. If his career is not, in the scheme of things, very important to his happiness overall, his sense of unhappiness with his career may decrease in intensity quite significantly. He may find, however, that this non-career-focused perspective no longer suits him. He may have adopted new career goals as a matter of personal fulfillment that weren't important to him before. If this is the case, shifting to a broader perspective that includes all of his personal goals may not move him to change his emotion after all.

v. Non-reflective emotion regulation strategies

Above, I focus on strategies for reflection upon our reasons. To adjudicate between emotions and judgments when they conflict with one another, I suggested using strategies that can mediate between conflicting emotions and judgments by allowing us to determine (1) what our reasons broadly are, and (2) how our reasons ought to be weighted given our total set of values and commitments.

Because emotions are maintained, in part, by feedback loops involving cognition, perception, and the body, there are also non-reflective strategies that we can use to regulate our emotions. First, there is peripheral feedback manipulation. For peripheral feedback manipulation, one modulates one's facial expression or posture in order to influence one's emotional state. Facial expression and posture produce bodily feelings. These bodily feelings may affect subsequent information processing by activating cognitive and perceptual processes that fit the profile of an emotional state that accords with the feelings one induces.[123] Second, there is situation selection or

[123] Philippot et al., "Cognitive Regulation of Emotion," 80-81.

modification. If one desires to modify one's emotional state, one can enter into a situation that would elicit a desired emotion, one can leave a situation that would elicit an undesired emotion, or one can modify one's situation so that only desired emotions will arise.[124]

The efficacy of such strategies might raise worries about the notion that we manage our emotions by reasons-reflection in a manner similar to our beliefs, intentions, and judgments. However, it should be noted that we have non-reflective means of managing these other attitudes as well. We engage in situation selection and modification to affect the sort of reasons we obtain for believing and intending. If one wants to believe something, one puts oneself in the way of evidence while avoiding situations that are likely to produce counterevidence. If one doesn't want to intend something one suspects one might, one avoids situations in which the intention would arise.

With our emotions, as with our other attitudes, we would be wise to wield non-reflective strategies for producing attitudes and attitude change with caution. It is better to reflect so that our attitudes will be guided by our reasons. If we don't reflect on our reasons, but implement a non-reflective strategy for eliciting a change in emotion, we may produce attitudes that are not apt to our situations or our total set of commitments and values. We may, at the same time, lose an opportunity to amend an errant judgment in accordance with the reasons that our emotions correctly took into account.

[124] James J. Gross, "Emotion Regulation: Conceptual Foundations," in Handbook of Emotion Regulation, edited by James J. Gross (New York: Guilford Press, 2007) 13-14.

Additionally, non-reflective strategies to produce emotional change are less likely to be effective if there really are no particular reasons to support the desired emotional state. Facial and posture feedback mechanisms facilitate emotions, and they appear to lower the threshold of evidence for emotion-corresponding features, but they don't guarantee that we'll enter into the emotion.[125] Ostensibly, if there are no facts or considerations at hand that could possibly support an emotional state, even with a low threshold for evidence, facial and posture feedback aren't likely to be effective.

Avoidance and distraction, while effective at keeping emotional states from arising for a time, do nothing to change our dispositions to experience them when we are forced to confront an emotionally charged object or situation. Moreover, avoidance and distraction strategies often backfire. Because we're engaging in guidance behaviors that stem from an intention to regulate our emotions, we reinforce the notion that the object or situation has the features that we wish not to confront. When the strategy backfires, our reactions to those objects or situations are prolonged and intensified rather than diminished.[126]

4. Conclusion

I began this chapter with a discussion of the relationship between our attitudes and our reasons. I claimed that although emotions can form automatically and be recalcitrant, emotions do have the sort of relationship with our reasons necessary to be

[125] See Furtack's review of feedback mechanism in Furtack, "Emotion, the Body, and the Cognitive," 54.
[126] See review of relevant research in Loewenstein, "Affective Regulation and Affective Forecasting," 184.

considered attitudes. I argued that attitudes need not be the products of an explicit reasoning process, though they are generally reason responding. Our emotions, I argued, meet the reason responding criterion. We can manage them according to our reasons by engaging in reflection. Emotions can be recalcitrant and thus fail to respond to reflection just as any of our attitudes can. However, emotions generally speaking are responsive to reflection on our reasons. Even in cases in which an emotion does not respond to judgment, judgment-oriented reasoning is just one modality of a broader reflective capacity. The emotion may respond to elements of our broader reflective capacity such as imagination or perspective taking.

Chapter 4

Helm's Dilemma Part 1: Irrationality without attitude conflicts?

In this chapter, I introduce Helm's dilemma, a dilemma that grew out of a debate about the nature of the irrationality of emotion/judgment conflicts.[127] The dilemma is this: Emotions either are attitudes or they are not. One the one hand, if emotions are attitudes, then emotion/judgment conflicts apparently are incomprehensibly irrational. On the other hand, if emotions are not attitudes, then emotion/judgment conflicts apparently are never even instances of ordinary irrationality.

The assumptions that create the dilemma seem plausible: emotion/judgment conflicts are irrational, and yet they are a *perfectly ordinary* form of irrationality. Since both our emotions and our judgments impact our thoughts, actions, and the plans that we construct, emotion/judgment conflicts have a disruptive quality that we're accustomed to viewing as irrational. They are also quite familiar. We don't experience them as rare and exceptional irrationalities, but as perfectly ordinary and commonplace. The dilemma is thus difficult to deny. If emotion/judgment conflicts are attitudes conflicts, they bear an uncomfortable similarity to judging A and not A simultaneously, which is a form of irrationality that is uncommon and quite perplexing. Such an irrational state is generally assumed to be beyond comprehension, if not impossible. We just don't know what it would mean to occupy a state of conflict so direct. Yet, if we adopt the view that emotion/judgment conflicts are not intentional attitude conflicts, it's not obvious that emotion/judgment conflicts are actually irrational at all.

[127] The dilemma is first introduced in Helm, *Emotional Reason*, 45.

Defenders of attitude views of emotions must take the first horn of the dilemma. They must show that the attitude conflicts involved in emotion/judgment conflicts aren't generally speaking incomprehensibly irrational. Defenders of non-attitude views of emotions must take the second horn of the dilemma, and show that an emotion that conflicts with a judgment typically violates some norm of rationality even though it is not implicated directly or immediately in an attitude conflict.

My treatment of the dilemma spans two chapters. In Helm's Dilemma Part 1, I survey the dilemma's origin and consider attempts to solve it by taking the second horn. Defenders of this option tend to emphasize the forceful character of emotions: emotions involve cognitive inclinations and motivations that dispose us to think and act in a manner that accords with our emotion even if we've refused assent to the emotion's representational content by judging that it's is false. This raises the possibility that emotions could be implicated in irrationality even if they are not attitudes. Because emotions dispose us to think and act in characteristic emotion-congruent ways, emotion/judgment conflicts might violate norms of rationality regarding coherence or the primacy of judgment.

In the chapter, I focus particularly on Michael Brady's view in his paper "The Irrationality of Recalcitrant Emotions."[128] In the paper, Brady proposes several norms of practical and epistemic rationality that our emotions allegedly violate when they conflict with our judgments. The norms he proposes are designed to show that it's irrational to be inclined toward attitudes that don't accord with what one has determined one's reasons to be in virtue of having made a judgment. Since emotion/judgment

[128] Brady, "The Irrationality of Recalcitrant Emotions," 413-430.

conflicts involve emotions that incline us to accept such non-judgment-concurrent attitudes, they are irrational on Brady's view.

Brady's view treats emotion/judgment conflicts as uniformly irrational in a way that implicates the emotions in falsehoods, rationalizations, and unreasoned attitudes. They are always, on his account, the source of the irrationality in an emotion/judgment conflict rather than the judgment. I suggest that there are two problems with this approach—one problem with the proposal itself, and one foundational problem with his approach to the issue.

The first problem with Brady's proposal is that his proposed norms are ill-founded. They are premised on the assumption that our judgment faculty has special authority in determining what our reasons are, and thus, what rationality requires of us. However, I argue that it is, in some cases, more rational to resist a judgment's directive than to conform to it. Sometimes our judgment faculty is vulnerable to errors that our other faculties avoid. Our conflicting attitudes and our emotions can, under the right circumstances, provide us with good reasons to revise our judgments rather than to defer to them.

The second, and more fundamental, problem with Brady's view is with his approach to resolving Helm's dilemma. His strategy is to show that emotions are always implicated in the irrationality of an emotion/judgment conflict in just the same way: they are the source of the rational conflict. This is, I contend, the wrong way to approach the problem. Conflicts between our emotions and our judgments don't all originate in our emotional faculties, which may be well functioning in some conflict cases. Our assessments of irrationality of an emotion/judgment conflict ought to reflect

the possibility that the irrationality of emotion/judgment conflicts might originate in a person's judgment faculty, in an erroneous judgment.

In Helm's Dilemma Part 2, I turn to the first horn of the dilemma. As a defender of a judgment view, Helm defends this option by emphasizing our apparent passivity with regard to our emotions.[129] When we experience emotion/judgment conflicts, we tend to experience our emotions as though the world draws them out of us, sometimes in spite of ourselves. He characterizes judgments, by contrast, as active. This distinction allows him to show that emotion/judgment conflicts are distinct from conflicts of judgment in a manner that demonstrates the former aren't as seriously irrational as the latter. What's unfathomable about two judgments with contrary content that are simultaneously assented to is the directness of the conflict. The conflict may be so direct that we can't imagine what it would mean for someone to endorse both contents sincerely and simultaneously. Thus, we might solve the dilemma by making a distinction between our more *active* judgment faculty and our more *passive* emotional states. If our judgments are active and our emotions passive, then emotion/judgment conflicts are more comprehensible. We can easily imagine actively endorsing one attitude while the force of habit causes us to passively harbor another, which should allay worries about inconceivability.

Though I agree with Helm that emotion/judgment conflicts are attitude conflicts, I disagree with his particular view. As with Brady's view, there is problem with the view itself, and a more fundamental problem with the approach—the very same fundamental problem that plagues Brady's view. The problem with Helm's view is that

[129] Helm, *Emotional Reason*, 45-6, 65-69.

the broad-sweeping claims he makes about emotions and judgments and their active and passive qualities don't reflect the whole range of emotions and judgments. It's moreover dubious that these claims hold in a sufficiently general manner to support relying on them to understand the nature of emotion/judgment conflicts as a general phenomenon. There are emotions that we actively cultivate and support, and there are judgments that come upon us rather passively. The more fundamental problem is that he treats emotion/judgment conflicts as if they were all implicated in the precisely the same variety of irrationality, and he treats emotions as the exclusive source of the irrationality involved. However, emotion/judgment conflicts are not all alike. They admit of important variations that impact the sort of irrationality that they involve.

In Part 2, I also present an alternative view of the dilemma and its solution that avoids the deeper problem with both Brady's and Helm's views. Helm's dilemma, as I interpret it, is essentially a Goldilocks problem. To satisfy the dilemma, we need to be able to show that on a spectrum from not-irrational-at-all to incomprehensible irrationality there is a middle ground that at least most emotion/judgment conflicts occupy. That middle ground is ordinary irrationality. To resolve the dilemma, it's not necessary to show that all emotion/judgment conflicts are irrational in precisely the same way, or that the source of the irrationality is always an errant emotion.

1. The dilemma's origin

Bennet Helm first issued the dilemma as a response to an argument levied against judgment views of emotion. In *Emotions and Reasons*, Patricia Greenspan argues that if emotions necessarily are, or involve, judgments or beliefs, then

emotion/judgment conflicts are logically incoherent.[130] The view that emotions and judgments are both attitudes involving assent to a proposition or representation of the world implies that emotion/judgment conflicts are, or are very close to, conflicts between contrary judgments. Conflicts between contrary judgments, it is generally supposed, are a form of irrationality so serious that they're hardly intelligible, if not psychologically impossible.[131] Intuitively, although emotion/judgment conflicts might very well be in violation of rational norms, they are not as radically irrational as this. On the contrary, they are perfectly ordinary and hardly unintelligible.

This argument is taken to count strongly against any view which suggests that emotions necessarily involve assent, as judgments do.[132] To avoid the implication that emotions necessarily involve assent, alternative views posit that emotions have representational content from which we can voluntarily withdraw or withhold assent, as do some other familiar mental states such as thoughts, construals, inclinations, or perceptions.[133] On these views, emotion/judgment conflicts arise when one's emotions persist in representing the world to be a certain way at the same time that one judges that the emotion's representational content isn't correct. For this reason, the experience of an emotion/judgment conflict is often compared to the experience of an optical

---

[130] Greenspan, *Emotions and Reasons*, 17.
[131] Aristotle humorously charges that to hold a contradiction is to be no better than a vegetable (1005b18ff). See Goldie, "Getting Feeling into Emotional Experience in the Right Way," 237.
[132] Sabine A. Döring "The Logic of Emotional Experience: Noninferentiality and the Problem of Conflict without Contradiction," *Emotion Review* 1.3 (2009): 240-241; Goldie, "Getting Feelings into Emotional Experience in the Right Way," 236-7; Brady, "The Irrationality of Recalcitrant Emotions," 415-416.
[133] The first three are cognitive non-attitude views; Perception views tend to be non-cognitive views.

illusion.[134] One can see a stick that is half-submerged in water as though it were bent without assenting to the perception that it really is bent. One can likewise experience fear without assenting to its representation of one's situation as dangerous.

In *Emotional Reason,* Helm argues that to deny that emotions necessarily involve assent invites a problem of its own. It's not obvious that the alternative non-assent, non-attitude views can meet their own expressed goal: to show that emotion/judgment conflicts are a perfectly ordinary form of irrationality. He remarks,

> Although these anti-judgmentalist accounts clearly avoid the problem of assimilating conflicts between judgments and emotions to incoherence, it is not clear that they are thereby able to provide a proper understanding of the nature of the resulting irrationality. After all, it is not at all irrational to have a stick half-submerged in water look bent even after one has judged that it is straight. If, as the anti-judgmentalist claims, the thoughts or construals emotions involve are not assented to, in what sense are they in rational conflict with judgment?"[135]

Thus, we are faced with a dilemma, which he articulates in the following passage:

> This seems to present us with a choice between two unsatisfactory alternatives: either we must be judgmentalists and accept an overly strong conception of rational conflict between emotions and judgment, or we must be anti-judgmentalists and give up hope of accounting for such conflict.[136]

The challenge of the dilemma is thus to do what each side of the debate has apparently failed to do: to show that emotion/judgment conflicts are irrational although they are not incomprehensibly irrational.

---

[134]Goldie, *The Emotions*, 74-6*;* Roberts, *Emotions*, 92; Tappolet, "Emotions, Perceptions, and Optical Illusions," 3-4.
[135] Helm, *Emotional Reason*, 43.
[136] Ibid., 45.

2. Motivation, inclination, and spurious reasons: cognitivist non-attitude views

To show that emotion/judgment conflicts might be irrational without being conflicts of judgment, proponents of cognitive non-attitude views typically capitalize on one very important difference between emotion/judgment conflicts and the optical illusions to which they are compared: a disavowed emotion can impinge on one's present thoughts in a way that a disavowed perception typically would not. Describing the difference between the experience of an optical illusion and the experience of fear, Roberts remarks,

> The fear has a personal depth and a life-disrupting motivational power that the illusion lacks. The bent stick is, at most, puzzling; the fear is personally compelling. This means that when the subject dissociates from his fear by denying its propositional content, it is like denying a part of himself, whereas denying his visual impression is not."[137]

Because emotions are connected to the person's values, Roberts argues, they continue to exert influence over a person's psychology in a way that perceptual illusions do not. They compel one to behave as though the emotion's representation were correct despite one's disavowal of the emotion's representation.

Peter Goldie portrays a similar picture of emotion/judgment conflicts when he suggests that emotions may "skew the epistemic landscape." Emotions, he argues, may have the effect of distorting our grasp of our reasons if they persist in opposition to our judgment. He writes,

> For example, we feel an extreme resentment at James, and we perceive grievous offence. But we believe that there really is no offence and the resentment is unjustified. However, rather than allowing internal conflict to arise and then to persist, and rather than abandoning our resentment, the mind seeks out all sorts of spurious "reasons" to support our resentment at James and our perceptions of his offence against us—"reasons" that would not come into view without the

[137] Roberts, *Emotions*, 92.

> emotional disposition being there in the first place: we come to believe that James has always held a secret grudge against us—and thus our resentment is justified; we come to believe that his apparent concern about our illness is really just a way of reminding us how much healthier he is than us—and thus our resentment is justified; and so on. In effect our resentment at James appears to us to be justified on the strength of spurious reasons that we have only arrived at *because of* our resentment.[138]

Our emotions have effects on our thoughts, the concepts we deploy in interpreting our situations, and our patterns of attention. These effects can draw out emotion-relevant considerations when we're assessing our situations, thus influencing our judgment. Goldie suggests that the reasons that our emotions draw attention to against our judgment are not really reasons for us at all by calling them "spurious."

Michael Brady, too, emphasizes the effects emotions have on our psychological state when we experience emotion/judgment conflicts. Because of the effects emotions have on our psychology, Brady observes that it may take greater focus and greater force of will for us to maintain attitudes that accord with our best judgment when we're experiencing an emotion/judgment conflict. The point can be illustrated with Goldie's resentment case from the above quote. Even if we've successfully resisted our inclinations to believe that James deserves our resentment, we may struggle to maintain this resistance as long as our resentment persists.

Given that emotions have these sorts of effects on our psychological states, proponents of cognitive non-attitude views have argued that emotions might violate, or cause a violation, of a norm of rationality. One possibility is that emotions cause attitude conflicts between our judgments and our beliefs, desires, and intentions. Recalcitrant emotions, Roberts emphasizes, have *motivational power* because they are

[138] Goldie, "Getting Feelings Into Emotional Experience in the Right Way," 238.

connected to the person's core concerns. For example, if you are experiencing what you judge to be inappropriate anger, you may retain a tendency to desire retribution or apologies. You may also retain a tendency to behave in a hostile manner toward the perceived perpetrator of the slight. Your dignity and social standing are important to you, so your continued anger will make it difficult for you to align your motivations to think, desire, and form intentions in a manner that coolly ignores your sense that you've been slighted. Thus, in such an instance, you might violate a norm of rationality by harboring incoherence between your judgments, desires, and intentions. Some of your intentions, desires, or other attitudes are in line with your judgment. Others are in line with a part of your psychology that you've repudiated even though, as Roberts put it, it remains "personally compelling."[139] These sorts of attitude conflicts represent a serious, but familiar form of irrationality. They are not as serious as conflicts between directly contrary judgments.

It's important to observe that, if this is the source of the irrationality of emotion/judgment conflicts, it's not obvious that the emotion in the conflict is implicated in the irrationality it causes. Emotions, on this sort of view, are at a level of remove from the attitude conflict that the irrationality apparently resides in. The emotion only puts one in the way of developing irrational attitude conflicts. Consequently, if one successfully manages to stave off the influence of an emotion on one's attitudes, it's not clear that there are any grounds for calling the emotion/judgment conflict irrational on its own. Accordingly, it isn't clear that there is any normative pressure to resolve it rather than to simply keep it at bay.

[139] Roberts, *Emotions*, 92.

Arguably, there is something irrational about unresolved emotion/judgment conflicts regardless of whether our struggle to avoid allowing an errant emotion to influence our attitudes is successful, and an adequate account of the irrationality of emotion/judgment conflicts should explain this. For this reason, Michael Brady argues that the irrationality that emotion/judgment conflicts represent might be better accounted for by another tack. On Brady's view, emotions don't necessarily result in the formation of attitudes that contradict the ones we've deliberated about and endorsed, but they do make it more difficult for us to have a sense of having settled the matter once we've formed a judgment. As Brady puts it, when we experience an emotion/judgment conflict, one remains *inclined* to accept the emotion's representation though one doesn't actually accept it.[140] To support his claim, Brady focuses on the cognitive effects of emotions, and in particular the effects of emotions on our patterns of attention. Our emotions first capture our attention, he argues, and then keep us consumed with the sorts of concerns that sustain the emotion. This primes us to assent to the emotion's representation. In cases of emotion/judgment conflict, we may not actually assent to the representation, but the cognitive effects of the emotion that prime us to assent to it are still operating.

According to Brady, having an inclination to accept an emotion's representation despite judging that the representation is false violates norms of practical and epistemic rationality on four counts:

(1) It is irrational to spend one's limited motivational energy and cognitive resources on matters that one judges unworthy of attention.
(2) It is irrational to be inclined to assent to a false representation.

[140] Brady, "The Irrationality of Recalcitrant Emotions," 421.

(3) It is irrational to be inclined to assent to a representation without good reason.
(4) It is irrational to be inclined to find or invent reasons to support a representation though there are none.

(1) is a norm of practical rationality. (2) – (4) are norms of epistemic rationality.[141]

Each of Brady's proposed norms are, I contend, ill-conceived. Each one assigns too much normative authority to our judgment faculty. To illustrate my arguments against Brady's norms, I borrow the following case from Catherine Elgin.[142] To paraphrase,

> Joan cares very deeply about her work in metaphysics and its reflection on herself. Consequently, her dispositions for envy track instances in which she is overlooked for her professional accomplishments in favor of peers who are, by her estimate, no more worthy of acknowledgement than she is. Her colleague Felix has just been named Mugwump Chair of Metaphysics and Joan is feeling envious. Via her emotion, she regards Felix as a recipient of valuable social and professional benefits or rewards that are not obviously deserved.

Joan's envy, we're told, has an effect on her patterns of attention. Her envy draws her attention to possible evidence of unfairness in the dispensation of benefits and rewards. The result is that Joan becomes aware of a pattern of discrepancies in the manner in which her work is received by her peers compared to Felix's: "Felix's manner is considered self-assured; Joan's is said to be aggressive. His first draft shows promise; hers needs work. His criticisms are incisive; hers are carping. His work extends the insights of his predecessors; hers are derivative." These differential terms, Elgin observes, "bring into focus the widespread practice of casting women's professional activities in a less favorable light than men's."[143] Joan's work and Felix's are being

---

[141] Ibid., 428-429.
[142] Catherine Z. Elgin, *Considered Judgment* (Princeton, Princeton University Press, 1996)149-150.
[143] Ibid.

assessed using different sets of concepts that reflect gender expectations. It becomes apparent to Joan that the possibility of a fair comparison between her work and Felix's is unlikely because she and Felix are judged by different measures.

Suppose that, in the end, Joan deems her resentful envy unreasonable. She judges that Felix is a good metaphysician with the sort of CV that would impress a selection committee for a prestigious position such as Mugwump chair. For her emotion to be justified, she supposes, she would need reasons to think that Felix wasn't truly deserving of special honors for his work and reasons to think that she might have deserved them just as well. Moreover, she judges that she really shouldn't be so concerned with recognition because she judges that the self-satisfaction that her work generates is what's really important to her. Since her envy is predicated on her concern for the value of professional recognition and accolades, she judges that it's inappropriate.

Suppose that Joan's judgment about Felix's desert and the inappropriateness of her envy does not dissolve her envy. Her envy persists and she experiences an emotion/judgment conflict. On Brady's view, she is irrational because her envy continues to dispose her to accept the emotion's representation of her situation in judgment. She is thus inclined to accept that which, according to her judgment, is a false representation regardless of whether she has what she judges to be good reasons to do so. She thus violates (2) and (3), which forbid inclinations to assent to false representations, and inclinations to assent without good reasons. She also continues to compare assessments of her work and Felix's though she has already judged that there is no relevant discrepancy to find with regard to the committee's decision to award

Felix the prestigious chair. She thus violates both (1) and (4), which enjoin continued attention to matters that one has settled for the sake of preserving cognitive and motivational resources, and to prevent the fabrication of reasons when there are none to be found. Since Joan' emotion drives her attention and her interpretation of what she observes, Brady's view suggests that her emotion-driven thoughts are mere rationalizations of her emotion; they do not provide her with reasons contrary to what she's already settled in judgment.

3. Against Brady's proposed norms

For reasons that will be familiar to readers from chapter 3, I hold that we ought to be guided by our reasons as a matter of practical and epistemic rationality, but I disagree with Brady's assumption that our judgments necessarily take precedence over our emotions when it comes to our reasons, practical or epistemic. We make mistakes in judgment with some regularity. Sometimes these mistakes are due to sloppy reflection on our reasons. Sometimes, we reason carefully, but form judgments on the basis of a too-narrow focus or sub-optimal weighting schemes. It isn't therefore safe to assume that our judgments are correct when they are at odds with our other attitudes, including our emotions. For similar reasons, I argue that Brady's proposed norms of rationality are ill-founded insofar as they rely on certain meta-normative assumptions about the primacy of judgment in showing that emotion/judgment conflicts are irrational.

Underlying Brady's norms, there appear to be two substantive meta-normative assumptions about our judgment faculty: An assumption that we ought to be ruled or guided by our judgment faculty, and an assumption that we ought to be resolute in our

judgments. The first assumption supports norms (2) – (4). The second assumption supports (1). I argue that both of these meta-normative assumptions are incorrect. The first assumption is misguided. The second is highly defeasible. Neither can justify the norms for emotions that Brady proposes.

i. The first meta-normative assumption

The assumption that we rationally ought to be ruled by our judgment faculty is widely taken for granted.[144] With our judgment faculty, we assess our reasons and form an opinion about which attitudes our reasons support. If a subject judges that p, p is what, by her estimation, her reasons support. It is, therefore, supposed that our judgment faculty has special authority in determining what our reasons require because it determines what is, by our own estimation, required of us. Consequently, if a subject judges p, but has attitudes that do not accord with p, she is irrational.[145]

Several of Brady's normative claims (2-4) align with this assumption. He suggests that if an emotion/judgment conflict occurs, our emotions ought to accord with our judgments.

---

[144] Notable dissenters include Harry G. Frankfurt, *The Importance of What We Care About: Philosophical Essays* (Cambridge: Cambridge University Press, 1988) 188-190; Alison McIntyre, "Is Akratic Action Always Irrational?" in *Identity, Character, and Morality: Essays in Moral Psychology*, edited by Owen J. Flanagan and Amélie Rorty (Cambridge, MA: MIT Press, 1990) 379; Robert Audi, "Weakness of Will and Rational Action," *Australasian Journal of* Philosophy 68.3 (1990): 279-281; and Arpaly, "On Acting Rationally Against One's Best Judgment," 491. See also Thomas E. Hill, Jr., "Weakness of Will and Character," *Philosophical Topics* 14.2 (1986): 93-115. Hill makes distinctions between cases of weakness of will that might shed doubt on the notion that all such cases are equally irrational.

[145] See, for instance, Scanlon, *What We Owe to Each Other*, 25; Michael Smith, *The Moral Problem* (Oxford: Blackwell, 1995) 248.

Brady's second proposed norm pertaining to emotion/judgment conflicts is that we shouldn't be inclined to assent to a false representation. Our emotion/judgment conflicts violate this, he suggests, because our emotions incline us to assent to a representation that is contrary to what we judge to be correct. Since Joan has judged that her envy is inappropriate, are we to presume that her continued envy does *in fact* contain a false representation of her situation? Suppose that she's misjudged her situation and her envy is, in fact, appropriate. According to Brady, regardless of whether one's judgment is *in fact* correct, it is correct *by one's own assessment*. This can have normative significance even in cases in which one has judged incorrectly. From your first person perspective, your judgment is correct and it's irrational to be inclined to hold contrary attitudes that you've deemed incorrect or inapt.

Brady's third and fourth proposed norms accord with Goldie's claim, noted above, that the reasons emotions bring to the fore are "spurious." His third norm is that we shouldn't be inclined to assent to a representation without good reasons. His fourth norm is that we shouldn't be inclined to find or invent reasons to support a representation though there are none. This suggests that our emotions can't provide us with good reasons contrary to our judgment if they provide us with reasons at all. Judgment takes precedence over our emotions in determining our best reasons.

The assumption that we ought to be ruled by our judgments is, I argue, misguided. Because we are prone to errors in judgment, sometimes even when we are deliberating carefully, we aren't normatively constrained by our judgment faculty in every case of internal conflict. Our norms should reflect the fallibility of our judgment faculty and the possibility of a need for revision.

Recall Nomy Arpaly's Susan *akrasia* case from chapter 3. Susan made a judgment that she ought to marry Todd on the basis of a too narrow subset of her reasons and using a suboptimal weighting scheme. As I described the case, Susan focused on a subset of her reasons that seemed most salient to her at the time: her desires to validate her current state of love and to secure financial stability, and her belief that marrying Todd would be the best way to satisfy these desires. However, she downplayed or ignored all of the evidence pertaining to other values that are truly more important to her such as respect and equal partnership. She does not expect that these latter values will be satisfied in a marriage with Todd. Her judgment may have seemed to her to be perfectly careful and balanced at the time she made it, but because it was in fact ill-conceived, it gave way to a desire not to go through with the wedding. Since Susan judges that she ought to marry Todd, her desire is akratic.

Note that a standard explanation of the irrationality involved in *akrasia* is that the akratic person acts on the basis of a subset of his reasons in spite of having judged that he ought to act otherwise on the basis of a more complete set of his reasons.[146] However, there are also cases of *akrasia*, like Susan's, in which the akratic person acts on account of a more complete set of her reasons than the one that formed the basis of her judgment.[147] Susan's desire not to marry Todd accords with the totality of her reasons better than her judgment does.

---

[146] Donald Davidson, "How is Weakness of the Will Possible?" in *Moral Concepts*, edited by Joel Feinberg (Oxford: Oxford University Press, 1970) 111; Christopher Peacocke, "Intention and Akrasia," in Essays on Davidson, edited by Bruce Vermazen & Merrill B. Hintikka (Oxford: Oxford University Press, 1985) 52.
[147] See McIntyre, "Is Acratic Action Always Irrational?" 388-389.

On the assumption that one ought to be ruled by one's judgment faculty, Susan is doubly irrational. She is irrational due to the attitude conflicts caused by an inapt judgment that she ought to marry Todd which doesn't accord with the totality of her reasons, and thus fails to accord with a number of her background beliefs, desires, intentions, and emotional dispositions. She also fails to form new intentions and desires to marry Todd in accordance with her inapt judgment's directive, making her akratic. However, the imperative to be ruled by her judgment, and thus to form intentions and desires that accord with it, seems dubious in this case. It would be worse overall for Susan to have a desire that is ill-justified on the basis of her total set of beliefs, desires, and values even though she judges that she ought to have it. She does better, overall, not to have it because it would conflict with her total set of attitudes just as her erroneous judgment does.[148]

It might be argued that even though Susan's *akrasia* works out fortuitously in this particular case, it is nevertheless a rationally defective way of preserving fidelity to our reasons. This becomes especially clear when we consider *akrasia* as a matter of character or disposition.[149] A rational agent whose attempts to guide her attitudes with judgment (or reflection more generally) are continually ineffectual is hardly an agent at all. Rational agency requires a high degree of effective control over one's attitudes, and failures of reflective guidance indicate a breakdown of such control.

[148] See Arpaly, "On Acting Rationally Against One's Best Judgment," 497-498.
[149] See Audi, "Weakness of Will and Rational Action," 280. See also Arpaly, "On Acting Rationally Against One's Best Judgment," 490. Arplay suggests that *akrasia* isn't something we could either deliberatively endorse or advise to others.

However, before we pass judgment on any particular case of *akrasia* and its bearing on a subject's rational agency, it's necessary to inquire into its causes. Susan's akrasia may very well originate in faculties that manifest her rational character—faculties over which she exercises control or influence.

Consider Harry Frankfurt's striking example:

> …suppose that someone made up his mind to let the world be destroyed for the sake of his finger, but that he could not go through with this intention because an overpowering surge of emotion prevented him despite himself from doing so. Suppose that his feelings revolted against his judgment, in other words, and that they constrained him from doing what he had deliberately made up his mind to do. The emergence into view of his incapacity to perform the action in question would substantially, or at least to some degree, *vindicate* his sanity. It would show that the irrationality displayed in his horrendous judgment did not, after all, go very deep.[150]

Frankfurt's view is that an inhibition to do what one judges one ought—which is often accomplished by "mobilizing [one's] emotions"—can be the result of a rational will. The will, according to Frankfurt, is no "featureless instrument" that we should expect to obsequiously follow the directives of our judgment. It has a content and character of its own, by virtue of which it places limits on what sorts of judgments we might be able to go along with. The character of these limits may be rational or irrational. Frankfurt suggests that one who is able to both form a judgment that world destruction is better than a scratched fingernail and to will accordingly exhibits worse irrationality than one who forms the judgment, but can't form desires or intentions that accord with it. The latter exhibits a rational will.[151]

---

[150] Frankfurt, *The Importance of What We Care About*, 189.
[151] Ibid.,186, 190.

Putting questions about the notion that *the will* is a substantial faculty that is itself rationally assessable aside, in the spirit of Frankfurt, we can ask whether the attitude-forming and regulating faculties that oppose a judgment might do so in a manner that manifests one's rationality. Karen Jones has argued that they may very well. According to Jones, as rational agents, we are committed to being guided by our reasons. To assume that what it is to be guided by our reasons is to be guided by our judgment faculty is to over-intellectualize this commitment, and to underestimate the force of this very same commitment with respect to our judgment faculty itself. Our commitment to being reason-guided is expressed in our commitment to (1) cultivating well-calibrated reason-tracking faculties, and (2) self-monitoring for errors. Supposing that we've kept these commitments with respect to the akratic attitudes that conflict with our judgments, on Jones's view there is no rational imperative to favor the judgment.[152] The akratic attitude may even have greater claim to our fidelity than the judgment in some cases. Our judgment faculty "can *fail* to express the commitment to rational guidance" if we fail to appropriately cultivate and monitor it for correctness in relevantly similar circumstances. She explains,

> On the proposed model, an all-things-considered judgment does not get normative authority for free. It has to earn such authority and it earns it in virtue of being the product of a conscious reasoning self that has itself been subject to regulation by reflective capacities."[153]

Pitted against a well-cultivated and well-monitored faculty, such a judgment has no special claim to rational authority.

---

[152] Karen Jones, "Emotion, Weakness of Will, and the Normative Conception of Agency," Royal Institute of Philosophy Supplement 52 (2003): 194-6.
[153] Ibid., 194.

Before we appraise Susan's failure to be guided by her judgment as irrational, then, it behooves us to ask whether her desires and intentions 'express her commitment to rational guidance,' to borrow Jones's manner of speaking. Even if Susan's judgment was carefully made, we know that it's short-sighted, and that further reflection on her total set of attitudes could have revealed its faults. Her judgment is thus defective as an expression of her commitment to rational guidance. Her judgment faculty isn't attuned to the relevant reasons and their relative weights, perhaps because she's failed to appropriately scrutinize her judgment faculty so that such errors could be avoided. Suppose, on the other hand, that Susan's desires and intentions tend to track her reasons accurately when she's making important life decisions. Suppose, also, that she has for some time monitored the deliverances of these faculties and, when she's detected errors, she's recalibrated them to prevent future errors. Her failure to have intentions and desires that comply with her judgments may therefore be an expression of her commitment to rational guidance. Susan is thus more rational in her akratic state than she would be were she to desire and intend according to her judgment's directive. In her akratic state, at least some of her attitudes are appropriate responses to her reasons.

In responding to conflicts between one's judgment and one's desires and intentions, then, one ought to adjudicate between them without prejudice in favor of one's judgment faculty unless one has good reasons to think that it is likely to be the best reason-responding faculty one has under the circumstances. If Susan realizes both that her initial judgment was poorly executed, and that her desires and intentions are generally reliable under the circumstances, it might even be rational for her to give some presumptive weight to the latter.

We can apply these same principles bearing on *akrasia* and rationality to emotion/judgment conflicts. In emotion/judgment conflicts, a bifurcation of correct judgment and incorrect judgment cases occurs that parallels the bifurcation of *akrasia* cases discussed above. In some emotion/judgment conflict cases, the emotion is ill-conceived because it arises in response to a too-narrow subset of one's reasons, while one's judgment faculty takes into account a wider set. In other cases, one's judgment is formed on a narrow set of reasons while one's emotion takes into account a wider set. Say Joan based her judgment that Felix was the deserving candidate for Mugwump chair on a too-narrow range of considerations. Her capacity for envy is well attuned to the fairness-relevant features of her situation. Her judgment, on the other hand, is clouded by a desire to remain in her colleagues' good grace in an environment in which suspicions of unfairness will be received as "sour grapes." When she forms her judgment, she thus focuses too narrowly on evidence that Felix is deserving and that the committee was fair, downplaying evidence to the contrary. Her persistent feelings of envy, in actuality, are more appropriate to her situation than her judgment because it is better supported by her complete set of beliefs, desires, and values. More and broader reflection on her reasons—reflection that her envy facilitates—could reveal that her judgment needs revision.

The imperative to change her emotion to accord with her judgment is thus, in Joan's case, poorly grounded. To have unqualified happiness for Felix would incline Joan to judge something that is, in fact, false: that Felix deserved the honor and that his having been chosen for it above other potentially deserving candidates (such as Joan, herself) was good and fair. By adjusting her emotion, Joan would only be adding fuel to

the conflict of attitudes divided by the narrow view of her reasons set before her in her faulty deliberations and the complete set of attitudes that she actually holds.

On Brady's view, Joan couldn't have good reasons to revise her judgment and retain her emotion even if her judgment was faulty. Her judgment determines what her reasons are, so if she deems her envy's representation to be false, she can't have good reasons to hold anything to the contrary. However, if Joan's envy is the result of an emotional faculty that expresses her commitment to rational guidance—she cultivates and monitors these faculties to ensure that they are well calibrated and highly accurate—her emotion may very well be supported by reasons pertinent to the matter of the Mugwump chair and its reflection on her wellbeing. These reasons, moreover, could justify an alteration of judgment.

Brady's norms (2)-(4) presuppose that our judgment faculty determines what rationality requires of us and that our emotions couldn't provide us with good reasons to revise our judgments. These norms are thus ill-conceived. To assume that we ought to be ruled by our judgment faculty is to afford too much rational authority to our judgments. We do sometimes have good reasons to revise a judgment on the basis of considerations that our emotions and other faculties bring to light.

ii. The second meta-normative assumption

The second assumption—that we rationally ought to be resolute in our judgments—is perhaps less widely held than the first assumption, but it has at least some *prima facie* appeal. To say that we ought to be resolute in our judgment is to say that we should not too readily re-evaluate our reasons or renege on a judgment once

we've committed to it. Suppose the impetus to re-evaluate one's reasons for a judgment is a desire or an intention that is contrary to it. We ought to be worried that the reasoning process that ensues will be riddled with ill-supported rationalizations masquerading as rational deliberations. Thus, assuming that one engaged in careful consideration of one's reasons in the formation of a judgment, a skeptical attitude toward one's inclinations to indulge in further deliberation may be called for. There are also practical costs associated with re-opening deliberations after forming a judgment. Even if we could ensure that deliberations undertaken in re-considering a judgment avoided rationalization traps, and could thus only improve the accuracy and appropriateness of our attitudes, our cognitive and motivational energies might be better employed elsewhere. Thus, re-entering deliberations can be a failure of procedural rationality.[154]

Richard Holton illustrates the rational failure of irresoluteness by describing an instance in which it's better not to revisit your judgment, even if you might correct an error by doing so:

> In general, once we have decided in which restaurant to eat, it is a good idea to let the matter rest, without endlessly discussing the pros and cons; and this is true, even though it occasionally means that we shall go to a restaurant that we could have realized was not the best choice.[155]

There are many considerations that one could weigh in determining which restaurant would be the best—the quality and quantity of the food options, the ambiance, the noise level, and so forth. It may not, however, be a good policy to continue deliberating about

---

[154] Richard Holton, "Intention and Weakness of Will," *The Journal of Philosophy* 96.5 (1999) 241; Michael E. Bratman, "Intention, Plans, and Practical Reason" (Cambridge, MA: Harvard University Press, 1987) 108-110.
[155] Holton, "Intention and Weakness of Will," 248.

where to eat once you've resolved to go to one in particular. Some reconsiderations rationally shouldn't be undertaken, though we may feel compelled to do so.

Brady's first norm is that one should not spend one's limited motivational energy and cognitive resources on matters that one judges unworthy of attention. According to Brady, when our emotions incline us to adopt an attitude that is contrary to our judgment, it causes us to attend to matters that we judge to be unworthy of attention, and to focus on reasons that we have, in judgment, deemed insufficient to support a contrary attitude. He suggests that our judgments determine what rationality requires from us in terms of our allocation of cognitive resources. This implies that we ought to be resolute in our judgments in emotion/judgment conflicts. Re-consideration of a judgment on the basis of an emotion is, he suggests, irrational. Presumably, he believes that any energy we devote to resolving the conflict ought to be aimed at managing our emotions so that they better align with our judgments and not vice versa. Joan's ends might be better served by focusing on other matters than whether Felix is better deserving of accolades than she is, or whether she, as a woman, is being judged by harsher standards. After all, the specter of rationalization looms for any further deliberation she engages in.

The assumption that we rationally ought to be resolute in our judgments is, I argue, attenuated by the possibility of a rational change of mind. A change of mind is not, in every case, irrationally motivated or impractical. The imperative to be resolute is therefore defeasible.

In her discussion of akrasia and the imperative to be resolute, Alison McIntyre suggests two ways in which the imperative to be resolute might give way to a justified

change of mind. First, new evidence might come to light. This new evidence could shift the balance of reasons in favor of a new attitude. Second, a contrary inclination points to the possibility that we've "underestimated the reason-giving force of the contrary inclination itself."[156] With Houlton's restaurant case, evidence that the restaurant is closed, or feelings of anxiety and regret that ensue once the decision has been made might justify revisiting the decision. If these feelings are appropriate responses to the situation, a change of mind about where to eat might justifiably be entertained. Suppose you are anxious about whether the restaurant can accommodate a member of your dinner party's severe food allergies. Assuming that you don't want your dinner companions to be harmed by your choice of restaurants, this is an important consideration that could potentially justify reopening deliberations.

Since Joan's envy persists in spite of her judgment, could she have sufficient reason to re-open deliberation about the fairness of her situation? Since her emotion tracks matters pertaining to fairness, perhaps she does. It's possible that she's misjudged the appropriateness of her envy. Further reflection could reveal that recognition really is important to her and that she has good reasons to want her work to be fairly judged by her peers. Moreover, due to the attention she paid to matters of fairness that her envy facilitated, she's apparently discovered evidence that she didn't have occasion to notice before: the pattern of apparent discrepancies in the manner in which her work and Felix's is judged.

[156] Alison McIntyre, "What is Wrong with Weakness of Will?" *The Journal of Philosophy* 103.6 (2006): 303.

It is of course possible that the attention one pays to possible new evidence or contrary inclinations in re-opening deliberation about a matter could be motivated in such a way that it distorts the truth and temps us away from a correct judgment. It is important to notice, however, that at no point in the process of forming an initial judgment and maintaining it is one immune to errors or rationalization. Most of the time we are not fully aware of the full range of influences and motivations that affect our judgment.[157] If Joan's judgment was marked with illicit motivations, calculation errors, or any other number of problems that were not obvious to her, her conflicting inclinations provide her with a clue that she might need to consider revising her judgment.

It may require a certain amount of self-knowledge and presence of mind to ensure that the adjudication of an emotion/judgment conflict or any other inclination/judgment conflict is conducted rationally. There are circumstances that flag the possibility of a judgment error that give us reason to re-open deliberation after we've settled on a judgment. We are especially prone to judgment errors in cases involving very complex practical calculations, for instance. In these cases, contrary inclinations may be sufficient evidence of the possibility of error to make re-opening deliberation rational. There are also reasons why we might suspect our emotions to be untrustworthy. If, for instance, we suspect the influence of phobias, depression, or various other disorders, it may be prudent to hold our inclinations against our judgment to be suspect. Once deliberation has been re-opened, care must be taken to reflect in a

[157] See Richard E. Nisbet and Timothy DeCamp Wilson, "Telling More than We Can Know: Verbal Reports on Mental Processes," *Psychological Review* 84.3 (1977): 233; and McIntyre, "Is Akratic Action Always Irrational?" 382.

manner that does justice to one's prior judgment. We ought to be wary of letting our inclinations carry us away. We can engage in strategies for attention redirection and information elaboration, described in chapter 3, to ensure that this is possible.

Assuming she has no particular reason to suspect that her envy is ill-calibrated and untrustworthy apart from its having conflicted with her judgment on this occasion, I contend that Joan is justified in reopening deliberation and reflecting on her reasons for her envy. Once she's reopened deliberation, she ought to engage in reflection carefully, making sure that her envy doesn't drive her judgment faculty without itself being subject to reflection.

The assumption that we ought to be resolute in our judgments is highly defeasible. We are vulnerable to errors in judgment that we can revise for principled reasons with the aid of our contrary attitudes and inclinations. The fact that our contrary attitudes and inclinations could cause us to rationalize or to have a skewed perception of our evidence gives us cause to be cautious about how and when we revise our judgments. It does not render it rational to ignore whatever reasons we might have for resisting the deliverances of our judgment faculty via our contrary attitudes or inclinations.

4. Conclusion

Brady's attempt to resolve Helm's dilemma by showing that emotions violate norms of rationality when they conflict with our judgments is based on dubious meta-normative assumptions that give our judgment faculty special authority in determining what our reasons are and thus what rationality requires of us. Norms of rationality based

on these assumptions have an unfortunate consequence. They suggest that it isn't rational to re-open deliberation or revise a judgment on the basis of an emotion. This is, I contend, wrongheaded. When our emotions are better supported by our reasons than our judgments, they can give us reasons to re-open deliberation and to change our judgments.

The more fundamental problem with Brady's approach is that the norms he proposes are targeted at showing that all emotion/judgment conflicts are irrational on account of the emotion, and that simply isn't the case. Our faulty judgments are sometimes the source of the irrationality, and an account of the irrationality of emotion/judgment conflicts ought to account for that. I develop this claim further in the companion chapter that follows.

Chapter 5

Helm's Dilemma Part 2: Attitude Conflicts and Ordinary Irrationality

In the previous chapter I presented Helm's dilemma, its origins, and its significance.[158] The dilemma is: Either emotions are attitudes or they are not. On the one hand, if emotions are attitudes, then emotion/judgment conflicts are apparently incomprehensibly irrational because they resemble similarly incomprehensible conflicts between two judgments. On the other hand, if emotions are not attitudes, then emotion/judgment conflicts apparently are never even instances of ordinary irrationality because there is no attitude conflict involved. Unless we are willing to give up the intuitively correct claim that emotion/judgment conflicts are irrational, but not extraordinarily so, then there is no easy escape from the dilemma.

In the previous chapter I also examined the possibility of resolving the dilemma by taking the second horn and showing that emotions could be implicated in irrationality even if they are not attitudes. I discussed Michael Brady's four proposed norms of rationality that emotion/judgment conflicts allegedly tend to violate.[159] I argued that Brady's proposed norms depend on dubious meta-normative assumptions about the rational primacy of judgment.

In this chapter, I consider the possibility of resolving the dilemma by taking the first horn, and showing that emotion/judgment conflicts could be attitude conflicts without being incomprehensibly irrational. As a defender of a judgment view, Helm has presented a view of emotion/judgment conflicts that aspires to do just that. Helm argues

[158] Helm, *Emotional Reason*, 45.
[159] Brady, "The Irrationality of Recalcitrant Emotions," 413-430.

that although emotions and judgments are both attitudes that involve assent, the character of the assent each involves is distinct. While judgments are an active form of assent, emotional assent is passive. This suggests that the character of emotion/judgment conflicts is different from the character of conflicts between contrary judgments. Although it may be difficult to conceive of actively and sincerely assenting to conflicting judgments simultaneously, it's not so difficult to conceive of passively assenting to something that you actively disavow.

In this chapter, I claim that Helm's view can explain how some portion of emotion/judgment conflicts escape extraordinary irrationality. However, I argue that the broad claims about emotion/judgment conflicts that the view assumes don't hold in such a general fashion that they can support a general account of the nature of emotion/judgment conflicts. There are more and less active and passive varieties of assent possible for both emotions and judgments.

Both Helm's and Brady's views, I contend, suffer because they approach the dilemma with the goal of showing that all emotion/judgment conflicts are irrational (but not extraordinarily so) in precisely the same way. They also both assume that all emotion/judgment conflicts are irrational in virtue of the emotion, rather than the judgment. On the alternative approach that I offer in the chapter, emotion/judgment conflicts fall on a spectrum of irrationality between not-irrational-at-all and extraordinarily irrational, and the source of the conflict—the attitude that requires revision to restore rational consistency—might be either the emotion or the judgment.

1. Active judgments and passive emotions: Helm's attitude view

In response to the dilemma, Helm argues that emotions involve a kind of assent that he calls "disclosive." Disclosive assent is implicit in one's evaluative feelings and in what he calls the *focal commitments* that one's evaluative feelings are indicative of. He characterizes evaluative feelings as feelings of pleasure or pain. These feelings demonstrate assent to the view that one's state of affairs is either good or bad. The focal commitments of one's emotions are commitments to *patterns* of emotions that accord with our attitudes about the import of focal objects. Say I value a statuette that adorns my living room. My statuette is a focal object. My focal commitments with regard to the statuette are commitments to experience emotions that correspond in tone to the status of my statuette. If my statuette is threatened, I experience anxiety or fear. If it's been broken, I experience sadness. I also have commitments to various forward and backward looking emotions that reflect this value as the situation changes over time. If I fear that my statuette will be broken as a result of your boisterous and cavalier behavior, I'm committed to being angry at you when you do break it.[160]

Disclosive assent characteristic of emotions is, according to Helm, "an acceptance that falls short of full blown judgment."[161] Disclosive assent falls short of a judgment, he explains, because it is passive.

> …they are states of consciousness that for the most part come over us in a way very much like that of perception, without our having to do anything more than (passively) be receptive to them. Indeed, this is the point of my claim that emotions are evaluative feelings: in feeling an emotion we are pleased or pained by the import of our situation impressing itself upon us.[162]

---

[160] Helm, *Emotional Reason*, 67-71.
[161] Ibid., 45-46.
[162] Ibid., 65.

Judgment, by contrast, is characterized as active:

> To make a judgment is to do something *actively*, consciously, and (for the most part) freely, and it is in large part for this reason that incoherence in judgment is unintelligible. That judgments are active, conscious and free in this way is part of what makes them so central to our cognitive lives.[163]

If emotions are passive in nature and judgments active, then emotion/judgment conflicts emerge in a manner that is arguably different from the manner that two contradictory judgments would. What makes an endorsement of two contradictory judgments so difficult to comprehend is that judgments are actively and consciously endorsed. We can't imagine what it would mean for someone to consciously and sincerely endorse two contradictory attitudes at the same time, without qualification. We can, however, imagine a person actively and consciously endorsing a judgment and, at the same time, passively assenting to some contradictory attitude passively through her emotions. We can imagine, for example, a person passively experiencing a strong sense of sadness that his prized statuette is broken, though he judges at the same time that the circumstances don't warrant any such sorrow. His emotion demonstrates a passive commitment to viewing the statuette as valuable and its destruction as regrettable, even though his judgment demonstrates conflicting actively endorsed commitments. The active/passive distinction, therefore, preserves our sense that emotion/judgment conflicts are irrational because they involve conflicting assent-bearing attitudes. Nevertheless, it does not implicate emotion/judgment conflicts in a kind of irrationality so serious that it's inconceivable.

[163] Ibid., 65.

Helm also uses the respective active and passive dimensions of judgments and emotions to explain why, on his view, emotions are generally the source of the rational conflict rather than judgments. He suggests that, as "active assents," judgments "normally have a kind of priority over emotions as articulations of our understanding of how things are." He goes on, "….given this priority, judgments normally have a kind of rational priority over emotions insofar as in cases of conflict between emotion and judgment it is emotion, as passive, that usually seems to thrust a vision of the world upon us contrary to what we in some sense really think."[164]

2. Against Helm's active/passive distinction

Our most passionate and forceful emotions (e.g. violent rage, devastating grief) are a stark contrast to our calm, protracted, and unhurried deliberations. We feel as though we have very little control over the former, though it seems as though we have quite a lot of control over the latter. When we experience strong emotions, they seem to come over us unbidden. When we form deliberative judgments, we calmly rehearse our reasons, make the relevant inferences, and come to conclusions as we see fit. Yet violent passions and calm deliberations hardly represent the most common occurrences of either our emotions or judgments. Most of our emotional experiences aren't particularly intense or imposing, and most of our judgments aren't calm or protracted. Most are more moderately situated somewhere between violent and calm, hurried and unhurried.

[164] Helm, *Emotional Reason*, 65.

To test Helm's account, then, I propose a look at the full range of emotion and judgment phenomena to see whether his characterizations of emotional and judgmental assent as passive and active, respectively, could hold in a sufficiently general fashion to support his view. This is the aim of the discussion that follows. I consider how Helm's judgment-descriptors, *active*, *conscious*, and *free* apply to a variety of judgments. I then examine how his emotion-descriptor, *passive,* applies to a range of emotions. I argue that, in large part, none of these descriptors significantly distinguish our judgments from our emotions across all, or even most, cases. The profile of the active or a passive dimensions of emotions and judgments as attitudes of assent can vary significantly from judgment to judgment and emotion to emotion.

i. Judgments

The *active* and *conscious* judgment-descriptors apply most readily to deliberative judgments that result from explicit reflection on one's reasons. Deliberation is a conscious process that requires directed attention and the application of reflective activities of a sort that might lead to a judgment. However, the *active* and *conscious* judgment-descriptors are less obviously or straightforwardly applicable to some other more spontaneous, non-deliberative varieties of judgment. Recent research on human mental processing suggests that the processes that issue in our judgments and other attitudes ought to be understood on a dual process framework. The dual process framework delineates processes that are fast and automatic from ones that are slow and effortful, requiring rule application and other sequential tasks, as well as the use of working memory. Members of the former set are called type 1 processes. Members of

the latter set are called type 2 processes. By comparison, type 2 processes require limited capacity and high effort. Type 1 processes require high capacity and low effort.[165]

Research on types 1 and 2 processes has revealed that, by comparison, we rely on type 2 processes quite rarely. The majority of our judgments issue from fast, automatic, and heuristic-driven type 1 processes which allow us to bring our background attitudes, skills, memories, and habits to bear on a matter. We do not rely on type 2 processes of conscious reflection or deliberation very often. Deliberation is slow, cumbersome, and cognitively costly because it requires the performance of sequential tasks in working memory. Such tasks take up a large portion of our cognitive capacity, preventing us from engaging effectively with other tasks that may require our attention. In studies, efforts to engage in effortful self-regulation and conscious reason-guidance resulted in poor performance on other simultaneous effortful tasks.[166] Additionally, such efforts produce lingering fatigue, so that even after a type 2 task is

---

[165] Each set of processes—type 1 and 2, respectively—is called System 1 and System 2 in much of the vast literature on the subject of the dual process theory of mind. There has been a recent movement away from referring to the sets of processes as unified systems however because there is evidence suggesting that the processes aren't integrated, for example because they can apparently conflict among themselves. On this matter, I follow Jonathan St. B. T. Evans, "How Many Dual-Process Theories Do We Need? One, Two, or Many?" in *In Two Minds: Dual Processes and Beyond*, edited by Jonathan St. B. T. Evans and Keith Frankish (Oxford: Oxford University Press, 2009) 34-5. See also Clare Saunders and David E. Over, "In Two Minds about Rationality?" in *In Two Minds: Dual Processes and Beyond*, edited by Jonathan St. B. T. Evans and Keith Frankish (Oxford: Oxford University Press, 2009) 317. Stanovich has, in a similar spirit, recommended referring to type 1 processes as "The Autonomous Set of Systems" (TASS) to emphasize their heterogeneity. See Keith Stanovich *The Robot's Rebellion* (Chicago, Chicago University Press, 2004) 37; and Stanovich, "Distinguishing the Reflective, Algorithmic, and Autonomous Minds," 56-57.

[166] John A. Bargh and Tanya L. Chartrand, "The Unbearable Automaticity of Being," *American Psychologist* 54.7 (1999): 464.

complete, one's ability to perform well on other effortful tasks is diminished. This effect is called "ego depletion."[167] We just don't have the mental energy or resources available to consciously guide our attitudes and actions very much of the time. When we do apply ourselves to the task of conscious self-guidance, we deplete limited resources that should be reserved for pressing matters.

Under the right conditions, reliance on faster automatic type 1 processes results in highly accurate, apt attitudes and actions, while avoiding the prohibitive cognitive costs associated with deliberation. This is especially important in circumstances that call for both quickness and precision. These advantages can be cultivated in the manner of a skill, or they can be innate. In performing a skilled activity, for example, playing a song on the piano, calculating intervals and finger configurations for each note would make the performance clumsy and slow. Hence, pianists typically practice their pieces *ad nauseum* until their skills become habits and their performances requires mostly sublimated routines needing little or no conscious thought or directed attention.[168]

An example of an apparently innate or unlearned type 1 process is the so-called "cheater detection" system that issues in fast, automatic judgments about who can or cannot be trusted. According to Cosmides and Tooby, success in social exchanges is facilitated by the ability to anticipate or detect cheaters—individuals who would receive benefits from us without paying the customary or agreed upon costs. On the surface, the reasoning involved in cheater detection could be domain general rule based reasoning.

---

[167] See Mark Muravan, Dianne M. Tice, and Roy F. Baumeister, "Self-Control as Limited Resource: Regulatory Depletion Patterns," *Journal of Personality and Social Psychology* 74.3 (1998): 774-776; Bargh and Chartrand, "The Unbearable Automaticity of Being," 464.
[168] Bargh and Chartrand, "The Unbearable Automaticity of Being," 468.

There is, however, evidence to support the view that humans evolved with a specialized domain-specific algorithm for processing information pertinent to cheater detection. This explains why test subjects frequently fail to exhibit understanding of the logic of conditionals on Wason selection tasks,[169] although they can quickly and correctly answer questions about how to determine whether people are following a rule of the general form: If you don't pay the cost, you can't have the benefit. This is evidence of an automatic domain-specific cognitive mechanism or ability that doesn't depend on the possession of domain-general reasoning abilities.[170]

Reliance on type 1 processes, whether they are cultivated or uncultivated, frees us up to engage in deliberation strategically, when a particular need arises, or when cognitive resources aren't better preserved for other tasks. Novel circumstances, error markers, attitude conflicts, and matters of importance are the sorts of things that might signal such a need, driving our decisions about when to be careful and methodical in our judgments rather than quick.

---

[169] In brief, the Wason selection task and its variations present test subjects with a material conditional statement. Participants are asked to check the veracity of the statement using a series of cards with only one side visible. The cards, which represent possible verifying or falsifying cases with which to test the statement have antecedents matching or non-matching symbols on one side and consequent matching or non-matching symbols on the other. Participants select which cards they believe would need checking. See P.C. Wason, "Reasoning about a Rule," *Quarterly Journal of Experimental Psychology* 20.3 (1968): 273-274. Details of the Wason selection task are further outlined in Chapter 6.

[170] See Leda Cosmides, "The Logic of Social Exchange: Has Natural Selection Shaped How Humans Reason? Studies with the Wason Selection Task" *Cognition* 31 (2009): 190-199. See also Leda Cosmides and John Tooby, "The Cognitive Neuroscience of Social Reasoning," in *The New Cognitive Neurosciences*, edited by Michael S. Gazzaniga (Cambridge, MA: MIT Press, 2000) 1260-1264.

Since our more spontaneous type 1 judgments don't proceed from a conscious series of reflections, the *active* and *conscious* descriptors aren't evidently appropriate, at least not in virtue of their having originated in an active and conscious deliberative process. Perhaps, in the case of non-deliberative judgments, those descriptors might fit our attitude monitoring and revising procedures instead. When it becomes clear that there is a particular need to prevent or correct for errors, slower, conscious deliberations can be implemented to replace or supplement our usually automatic procedures. Monitoring and revising routines, however, need not be consciously implemented. They, too, may be automatically activated and conducted, requiring no conscious attention or choice.[171]

For example, one study identified college students who reported having a goal to perform well academically which they associated with a desire to please their mothers. These same students performed better on a test if they were subliminally primed to think about their mothers than students who either weren't primed or had previously reported no academic achievement goals associated with their mothers. Accordingly, Fitzsimons and Bargh argue that thinking of persons with whom we share goals can result in automatic, unconscious goal activation, resulting in the activation of an automatic self-regulation routine. Persons who we associate with interpersonal goals are just one of many possible external triggers of unconscious goals that results in automatic regulation of attitudes as well as behaviors and emotions.[172] The *active* and

[171] Fitzsimons and Bargh, "Automatic Self-Regulation," 152.
[172] Gráinne M. Fitzsimons and John A. Bargh, "Thinking of You: Nonconscious Pursuit of Interpersonal Goals Associated with Relationship Partners," *Journal of Personality and Social Psychology* 84.1 (2003): 148-164; Fitzsimons and Bargh, "Automatic Self-Regulation," 158.

conscious judgment-descriptors that Helm invokes may thus apply to only a limited portion of our judgments that involve either deliberation or self-directed monitoring and regulating.

The *free* judgment-descriptor also needs some qualifications for both deliberative and non-deliberative judgments. When it comes to assent-bearing attitudes such as beliefs or judgments, we lack perfect control. We do not generally speaking have the power to judge just anything it might please us to (e.g. 'unicorns exist;' 'I am a wizard alumnus of Hogwarts School') irrespective of the reasons that we have or lack that could either support or rule out such an attitude. Our judgments and beliefs are guided and constrained by the force of reasons that we are faced with that are convincing to us in one way or another.[173] Our freedom with respect to what we believe or judge is limited to that which has, for us, some sort of truth-conceit. We can't have whatever assent-bearing attitudes we wish in earnest because such attitudes are the sorts of things that purport to represent reality. We must, in some sense, regard them to be true.

Even though our judgments are so-constrained, we do exercise some measure of control over them insofar as we are capable of influencing them in each of the following three manners. First, we can direct our attention to a particular attitude's content and evaluate the reasons we has that support having, maintaining, or extinguishing the attitude. For example, if I have judged that a particular job candidate is the best one for

[173] Bernard Williams argues that beliefs are necessarily constrained due to the fact that they purport to represent truth. See Williams, *Problems of the Self*, 148. See also Mark Thomas Walker, "The Voluntariness of Judgment," *Inquiry* 39.1 (1996): 99-102. Walker offers a helpful discussion of the significance of the truth aim of judgment for the voluntariness of judgment.

the job I'm hiring for, I can, if I choose, reflect on my reasons for and against my judgment as often as I please to ensure that my judgment is well-founded. Second, we can engage in further inquiry in order to seek out reasons bearing on the attitude. For example, I can do more research on all of the job-seeking candidates under consideration to see whether there are pertinent features of their record that I overlooked or under-weighted before. Third, we can engage in activities aimed at changing states of affairs so that we will have the reasons to judge that we wish to have. If I want to believe that my favored candidate has been selected, I can use whatever social capital I have to ensure that she is.[174]

The extent to which we are able to successfully exert control over the individual judgments that we hold can be more or less robust. Most obviously, my ability to alter states of affairs to fit what I want to believe or judge extends only to that which is in my power to bring about. If I have no power to ensure that the hiring committee will select my favored candidate, I can't bring myself to believe that he or she is hired by making it so. Additionally, suppose that the negation of a belief or judgment may be so unthinkable or upsetting to me that I'm unable to will myself to re-evaluate it or inquire into it any further. This would undermine my ability to exert control over that particular attitude. I won't be able to influence it by engaging in reflection or inquiry. Finally, my beliefs and judgments may be recalcitrant. They may not be appropriately responsive to

[174] The three forms of control I discuss here are informed by Hieronymi's discussion of control and self-management. Hieronymi distinguishes two varieties of control: evaluative and managerial. Evaluative control is exercised by considering one's reasons. Managerial control is exercised by seeking out reasons or manipulating states of affairs. See Hieronymi, "Controlling Attitudes," 53, 55-56.

whatever reasons I may have in favor of modifying or extinguishing them. In the most extreme cases, they may be wholly impervious to the influence of reasons.

In summary, Helm's judgment-descriptors, *active*, *conscious*, and *free* don't fit all, or even most, of our judgment phenomena in the same way or to the same degree. Deliberative judgments involve reflective activities that are active and conscious, but the majority of our judgments aren't arrived at by deliberation. They are spontaneous and automatic. These sorts of judgments might be actively and consciously monitored and regulated, but they usually are not. Our monitoring and regulating processes tend to be automatically initiated and implemented as well. Thus, the *active* and *conscious* descriptors seem to fit only some of our judgments. The applicability of the *free* judgment-descriptor, I suggested, depends on our ability to exert control over our judgments. I claimed that we have control over our attitudes insofar as we are able to affect them by re-evaluating our reasons, inquiring in order to expand our reasons-set, or changing the circumstances to reflect the reasons that we wish to have. The amount of control that we have over any particular attitude may be more or less perfect.

ii. Emotions

The *passive* emotion-descriptor applies most readily to passionate emotions such as violent rage and devastating grief. When we experience these sorts of emotions, it's usually because we find ourselves confronted with situations with deeply important value relevant features. Our emotional responses in these instances may arise without hesitation or precipitating thought, giving them the appearance of having arisen unbidden. The thoughts, feelings, and behavioral tendencies that attend them are vividly

experienced as the significance of the situation impresses itself upon us, and we may have little opportunity to interfere with their occurrence until after they've already begun to take effect. The stronger the emotion, the more difficult interrupting the emotion's progress may be. When one is experiencing fresh grief, for example, the inclination to cry or collapse, or to dwell on the dreadfulness of the loss one has suffered might be overwhelming. It may take concerted and persistent efforts to stay these impulses should the necessity arise. However, most of our emotional experiences are not quite so volatile and forceful as our most powerful rage or grief.

These particularly intense emotions tend to accompany gains, losses, threats, and so forth that are of *extraordinary* and *immediate* importance. Unless you are in peculiarly unfortunate or unstable circumstances, most of your emotions will be responses to more mundane stimuli, which typically elicit more moderate emotions. Our more moderate emotions aren't quite so intense or forceful in our experience, though they may still arise spontaneously, without summoning.

The suddenness of an emotion's onset may make it appear as though it is out of one's control because it arises without a precipitating conscious or self-aware endorsement. As Robert Solomon argues, the suddenness of the situation that one find's oneself in contributes to the emotion's reflex-like appearance. Nevertheless, an emotion is typically a genuine *response* to one's situation, not a mere reflex. He provides a useful example in the following passage:

> I am driving along a mountain pass and I suddenly see a rock slide in front of me. The rock slide is unexpected and does indeed "happen to me," but my emotional response is quite a different matter. Depending on my driving skills, my self-confidence, and my previous experience (not to mention my temperament, tendency to panic, etc.), both my emotional response and my actions (which cannot be easily separated) are just that, *my responses*. They may

> be spontaneous, unthinking, and, if I am practiced in the art of driving in dangerous conditions, habitual. My response need not be fully conscious. It certainly need not be articulated or explicitly "thought" at the time. There is no room for deliberation. What I do and feel no doubt depends on my history of habits and kindred experiences, but it is the situation, *not my emotion*, that suddenly confronts me.[175]

Even though an emotion may arise without thought or hesitation, emotions are very different from knee jerks, eye twitches, and other mere reflexes. Reflexes arise unbidden and are wholly unconnected to any normative reasons or commitments. On the other hand, our background attitudes, perceptual skills, memories, expectations, and habits are expressed through our emotions as we monitor our environments (consciously or not) for matters of personal importance.

Our spontaneous and automatic emotional states are also distinguished from reflex responses by our ability to maintain or modify them with reflection on our reasons. Unless an emotion has become unusually recalcitrant, we can modify it to better reflect our reasons using strategies that employ imagination, information elaboration, and attention redirection. As I argued in chapter 3, these reflective activities draw out reasons in a manner that engages our emotional faculties, even when other forms of reflection might have failed to do so.

The *passive* descriptor is least obviously applicable to emotions that emerge from the activities of reflection. Particularly when we're encountering novel situations, we may not have a strong automatic emotional response other than, perhaps, curiosity. When we don't know how to feel, we can engage in reflection on the situation's value relevant features in order to elicit an appropriate emotional response. Once we are in an

[175] Solomon, *Not Passion's Slave*, 201.

emotional state, or we have formed an emotional disposition, we can re-evaluate the reasons that support having it or extinguishing it with reflection. We can also change our situation in order to change the reasons that we have, and we can seek new reasons to supplement the ones that we already have by engaging in further inquiry. Our attempts to affect our emotions may not have perfect or immediate success in any particular instance. However, we generally can exert some control or influence over our emotions, both by actively and consciously reflecting on our reasons and by establishing new emotional dispositions and automatic regulation routines.

To recap, emotions have a passive appearance in virtue of their sometimes spontaneous and automatic origination. Nevertheless, emotions admit of more active varieties and features, too. We can use reflection to establish, extinguish, or alter our present emotions or emotional dispositions.

iii. The active/passive distinction

I am now in a position to assess whether descriptors such as *active*, *conscious*, and *free* really distinguish our judgments from our *passive* emotions in a way that warrants Helm's use of these distinctions in a general account of emotion/judgment conflicts. The senses in which we might be generally *active*, *conscious*, and *free* with respect to some of our judgments, I will argue, don't necessarily distinguish them from emotions in a significant or sufficiently general manner to support Helm's account.

Both our emotions and our judgments can arise from or be modified in response to conscious reflective activities such as deliberation, and both can be subject to active and conscious monitoring and regulating. Thus, our judgments and our emotions both

admit of active and conscious varieties. On the other hand, both our emotions and our judgments can arise automatically, without the guidance of deliberation or any other reflective activity. When they do arise automatically, we might engage in active and conscious monitoring and regulating of our judgments or our emotions, but these processes can be automatic as well. [176] Our judgments and our emotions both therefore admit of passive varieties.

The sense in which we are *free* with respect to our judgments also does little to distinguish them from our emotions. The freedom that we have with respect to our judgments is in our ability tocontrol or influence our judgments in particular ways. We can't bring ourselves to have just any judgment we please. We are guided and constrained by what we can earnestly assent to. Our emotions are similarly constrained. I cannot, for example, be afraid of bunny rabbits in earnest just because I might fancy doing so unless I can manufacture some reasons that would make it possible for me to view bunnies as a real, compelling danger. I cannot be wholeheartedly happy about the death of a dearly loved one, no matter how much I wish to make the pain of grief end, unless I can find reasons to view the loss to be an unqualified good for me. Our emotions, like our beliefs and judgments, are reason responding. They are generally the earnest reactions of our emotional faculties to what we've perceived to be the case.

Although our judgement faculty is not without constraints, we can exercise control over our judgments in a number of ways, including reflecting on our reasons, seeking additional reasons, and manipulating our circumstances. We can do the same

[176] See Fitzsimons and Bargh, "Automatic Self-Regulation," 152. Fitzsimons and Bargh's research suggests that automatic, unconscious goal activation is involved in judgment, emotion, and behavior regulation alike.

with our emotions. For instance, if my colleague says something particularly cutting and rude at a meeting and I'm angry about it, I can re-evaluate all of the relevant considerations that support or contravene my anger if I choose to do so. I can also seek new evidence that I don't already have regarding the situation. For instance, if my mind was wandering during the meeting, I might inquire into the possibility that I heard the comment out of its intended context. Re-evaluating the situation, with or without new considerations to add, might result in changes in my anger, or in its dissipation. Additionally, if I want to have reasons to be happy with my rude colleague at future meetings, I can approach him and talk through our differences with the aim of repairing the relationship.

Our attempts to influence any particular judgment might be more or less successful, so our control over our judgments might be less than perfect. The same is true of our emotions. We of course don't have perfect control over just every external circumstance that our emotions might reflect. For example, my colleague might be unwilling to agree to be civil, which prevents me from changing the situation so that I'll have the reasons to be happy that I want. Some of our emotions may be stubbornly difficult to move. We also may sometimes harbor wholly recalcitrant emotions. If an emotion is recalcitrant, our attempts to influence it by reflection on our reasons won't be successful because recalcitrant attitudes don't respond to reasons. Our emotions and our judgments both admit of cases that range from the easily influenced to the immovable.

My conclusion is that the active judgment/passive emotion distinction fails to hold across the full range of emotion and judgment phenomena. There are emotions

with active features: we engage in reflection to arrive at or influence them. There are also judgments with passive features: they are automatic, and they can be difficult (or, in extreme cases, impossible) to control or influence.

The failure of the active/passive distinction to hold all-inclusively or nearly so is a serious problem for Helm's view. What makes a conflict of contrary judgments incomprehensible, according to Helm, is that we can't imagine a person sincerely and knowingly giving her active assent to representations with inconsistent or contradictory content at the same time. We can, however, imagine someone actively endorsing a representation via a judgment while passively assenting to a contrary representation via an emotion that conflicts with it at the same time. Therefore, assuming that the active judgment/passive emotion distinction holds quite generally or universally, Helm can show that emotion/judgment conflicts are irrational, though not incomprehensibly so, even if they are contrary assent-bearing attitudes. But since the distinction doesn't hold generally, his account can at best show that conflicts between emotions and judgments that reflect the active judgment/passive emotion characterizations meet the parameters that Helm's dilemma sets on accounts of emotion/judgment conflicts.

It might help Helm's account to establish that our degree of control over our emotions is typically less perfect than our control over our judgments. Perhaps, on average, it takes greater or more protracted, incremental efforts to influence our emotional responses. However, it's not obvious that this is so. If it seems obviously true, it may be because we're accustomed to thinking in terms of the skewed stereotypes I described at the beginning of section 2: our most passionate and violent emotions stand as our representatives of the whole domain of emotions and our most measured

judgments stand as our representatives of the whole domain of judgments. However, these stereotypical emotions and judgments are hardly the mean. They actually make up only a small portion of the domains they're supposed to represent. When we shift our focus in the direction of more ordinary cases rather than these rather extreme ones, the differences between our emotions and judgments no longer appear to be so great. Our beliefs and judgments can be quite deeply rooted and difficult to move, just as our emotions can. By contrast, our emotional responses might change quite readily in response to newly presented reasons if we are attending to them carefully.

Moreover, even if it is generally true that on average our emotions are more difficult to influence than our judgments, the difference in degree of control across all instances is not enough to support Helm's general account of emotion/judgment conflict, which depends on the truth of the active/passive distinction. The active/passive distinction doesn't hold across the full range of emotion and judgment phenomena. Some conflicts will involve outlier emotions and judgments: active emotions that are arrived at by means of reflection and are easily controlled, and passive judgments that arise spontaneously and automatically and are recalcitrant. An adequate account of emotion/judgment conflict should acknowledge the full range of conflict possibilities involving active and passive varieties of both emotions and judgments. I offer such an account in section 3.

3. Toward a better account

I begin my account with a summary of the factors that contribute to the diversity of emotion/judgment conflict cases. Some of our emotion/judgment conflicts are, just as

Helm suggests, conflicts between an active sort of judgment and a passive sort of emotion. Others involve a variety of active and passive dimensions which are reflected in the extent to which an attitude is either automatic or actively and reflectively endorsed, and the extent to which we are capable of controlling or influencing an attitude with reflection. I suggest that there are three broad conflict possibilities with many variations: passive emotion/active judgment conflicts; active emotion/passive judgment conflicts; and passive emotion/passive judgment conflicts. An additional possibility is: active emotion/active judgment conflicts. It may be that at least some active emotion/active judgment conflicts would be so direct, and thus so inconceivable, that we might conjecture that they are beyond the threshold of possibilities on the spectrum. If such extreme cases are possible, they aren't a part of our ordinary experience of emotion/judgment conflicts, which tend to fall in the three categories listed above. We can think of active emotion/active judgment conflicts as limiting cases.

Because there are differences in the kinds of emotion/judgment conflicts that are possible, there are also differences in the kind and severity of the irrationality that they represent. An emotion/judgment conflict can be more or less direct. The more direct the conflict, the more irrational it is. A conflict can be direct to a greater or lesser extent in virtue of two aspects. First, it can be direct in virtue of the *attitudes* involved; Active assent to contrary attitudes is more irrational than active assent/passive assent conflict. Second, a conflict can be more or less direct in virtue of the *content* of the attitudes involved; To judge *A* and *not A* is more direct than judging *A* and *B*, though one also holds conflicting attitudes with the content '*B* entails (or, less strongly, supports or probabilifies) *not A*.'

Put together, the variety of emotion/judgment conflict cases suggest that there is a spectrum that these cases fall on. Individual conflicts might occupy a number of points on this spectrum, with rationality on one end of the spectrum and extraordinary irrationality on the other end. I'm suggesting that most cases of emotion/judgment conflict fall somewhere between mildly irrational and extraordinarily irrational. To illustrate, I examine examples of each of the three broad possibilities mentioned above of conflicts that vary on the active and passive dimensions of the attitudes involved in a particular conflict. I also discuss a fourth limiting case. With each example, I examine how the directness and control factors influence where on the spectrum a case falls and what kind of normative pressure there might be to restore rational consistency.

i. The limiting case

I begin with the limiting case: an active judgment/active emotion conflict. It is the limiting case because it is the sort of conflict that goes beyond what we can fathom an ordinary agent to be capable of. For example, Say that Frank isn't really angry at his coworkers for dismissing or ignoring his ideas in a meeting, and on reflection he doesn't think he has reasons to support a judgment that he's being slighted. However, he suspects that a genuine display of anger will get his colleagues' attention and command respect for his ideas, due to the passionate endorsement that his anger will convey. The sort of respect for his ideas that he might generate this way might also have the effect of generating greater social capital that he could wield to his advantage in future meetings, and in advancing his career. Perhaps consciously, or perhaps not, Frank

wants to be angry at his coworkers for behaviors that are, he judges, too small to justify anything stronger than mildest annoyance.

Being an ordinary human agent, being angry isn't something Frank can simply decide to do. He will have to draw an anger response out of himself by a number of intermediary actions: changing his attention patterns, his focus, and his conceptualizations of the situation. To draw out anger, Frank focuses on any body language of his colleagues' that could possibly read as dismissive or contemptuous. He pays particular attention to any comments that could be interpreted as irreverent, so that he can dwell on the possibility that he's been slighted. Then, when a colleague interrupts him in the middle of a sentence to ask a pointed question, it happens: anger wells up.

In actively drawing out anger, Frank suppresses his once active judgment by changing the way that he presently views his reasons. Either he takes the judgment offline and compartmentalizes it, or else he changes his mind. Otherwise, it's highly doubtful that his attempts to make himself angry could be successful.

Active judgment/active emotion conflicts, are the limiting case on my account because, as Frank's case illustrates, for ordinary agents, actively and reflectively cultivating an attitude *A* requires either removing assent to *not A* or removing attention from the sorts of reasons that support *not A.* Actively cultivating and assenting to A and not A attitudes isn't apparently something we can do simultaneously. This is what makes judgment/active emotion conflicts unfathomable. They are beyond the sorts of irrationality that are part of our ordinary experience, and they are beyond the sorts of irrationality that we can imagine other ordinary agents sustaining.

There might be some variations on the active judgment/active emotion conflict case that would turn out to be within the realm of possibility because other dimensions of the conflict affect the irrationality of the case. For example, imagine that Frank's judgment about his colleagues is only indirectly in conflict with his anger in virtue of the attitudes' content. Instead of judging that there is nothing to be angry about in this instance he holds a more general judgment that his colleagues are agreeable, cooperative, well-intentioned, and polite. This is a sort of judgment which might be consistent with the attitude that his present situation is an exception. To the extent that he has reasons to view it as such might increase the likelihood that he would be able to actively cultivate and reflectively endorse both his more general judgment and his anger in the present case.

ii. An active judgment/passive emotion conflict

An active judgment/passive emotion conflict involves a reflective sort of judgment that is subjected to monitoring and is sensitive to reasons. It also involves an emotion that arises automatically and isn't particular well-monitored or conscientiously cultivated. For example, Oliver has decided to travel by airplane for the very first time. He has done his research and has satisfied himself that airplanes are, generally speaking, a safe means of travel. He recognizes that airplanes do crash on rare occasion, but that crashes are a very low risk, statistically. Being new to the experience of flying, however, Oliver doesn't quite know what turbulence feels like, though he's read about it. On his flight, there is some fairly strong turbulence during a thunderstorm and he

experiences some fear and anxiety. He worries that the plane is going to crash, even though he hasn't changed his mind about his previous judgment.

Oliver's emotional reaction is passive. It is wholly untutored and automatic. He doesn't have any sense of what normal turbulence is like to allow him to gauge what sorts of movements of the plane might be indicators of genuine trouble, so his fear is excessively sensitive to the jerks and wobbles that he feels on his flight. After all, jerks and wobbles of the sort he's experiencing would be justifiably worrying on other modes of transport. His judgment, on the other hand, has been meticulously shaped by research and careful reflection, and he understands that turbulence is a normal and generally non-threatening occurrence. He doesn't have any particular reason to think that his emotional response is likely to track actual danger in what is, to him, a novel situation.

Oliver's emotion/judgment conflict is irrational. It's a conflict that will disrupt his normally more unified and cohesive way of thinking and guiding his actions. It nevertheless falls short of extraordinary irrationality because his *attitudes* don't clash directly, in the manner of two actively cultivated or endorsed but inconsistent attitudes. His judgment is actively endorsed, but his emotion is comparatively quite passively borne. Without this aspect of directness, the conflict would not fall beyond an acceptable range on the spectrum of cases.

Notice, also, that in Oliver's case, as I have described it, there may or may not be a very direct conflict in virtue of the *contents* of the attitudes involved. His judgment that airplanes are highly unlikely to crash may conflict with the fear he experiences on an airplane, but a judgment such as Oliver's need not conflict with fear in every possible flying scenario. If he has reasons to support viewing his situation as one of

those possible, though statistically unlikely, crash scenarios that his judgment can accommodate, then his attitudes are consistent. The irrationality of his emotion/judgment conflict is reduced to the extent that he has reasons for viewing his situation as one of those exceptions. If he has very good reasons to view his situation as an exception, then his emotion/judgment conflict isn't obviously inconsistent or irrational at all.

iii. A passive judgment/active emotion conflict

An active emotion/passive judgment conflict is one which typically involves an emotion that is conscientiously and reflectively monitored or maintained. It may even be bolstered by reflection on one's present situation and the reasons it proffers in support of maintaining the emotion. An active emotion/passive judgment conflict will also involve a judgment that is automatic. For example, Doris is an experienced businesswoman with well-honed intuitions about when she is being swindled.[177] Over the years, she has learned to pick up on subtle cues regarding a contact's transparency and follow-through. Evidence of evasiveness or dissembling with furtive actions or attempts to quell worries with ambitious promises are, in Doris's experience, reasonably good evidence of untrustworthiness. Her sense of suspicion generally tracks just these sorts of cues, and she closely monitors her suspicion for potential misfires by reflecting on her reasons and her past track record.

[177] This case is adapted from one Greenpan discusses. See Greenspan, *Emotions and Reasons*, 3-4. Greenspan reprises the case again as an extended example for chapter four of the same book.

Doris has a mild sense of fear or suspicion about her business contact Floyd. Floyd has triggered her emotion because he seems ever-so-slightly oily. He appears to be particularly keen to hook her on promises that she finds to be just one small step beyond what she'd usually find to be reasonable. On reflection, Doris notices that Floyd has these features that normally justify her suspicion, even if they are only subtly present in the particular case. However, Floyd also has all the sorts of features that generally trigger automatic judgments of trustworthiness. He is a tall Caucasian family man with a well-trimmed beard, he speaks with a generic Midwestern American accent, and he is a generally well-liked member of the community. Doris is inclined to trust people like him, and her automatic, non-deliberative judgment is that he is very likely to be worthy of her trust.

On a cursory sort of reflection, she wonders whether her emotion is a manifestation of her prejudice against young business people who might seem just slightly overeager. After more careful reflection, she might find either that her suspicion is justified, or that her emotional faculties have misfired in this instance. If her only evidence of his trustworthiness is a set of surface features that don't directly bear on his trustworthiness such as his race, physical appearance, and family status, then her reasons for her emotion are likely better ones than the ones supporting her judgment.

The *contents* of Doris's attitudes are in direct conflict. In her fear and suspicion, she regards Floyd to be untrustworthy at the same time that she judges that he is trustworthy. The *attitudes* in conflict however, don't have an equal share of active qualities. Doris' emotion is actively maintained, and it apparently stands up under reflection on the sorts of reasons that normally support her suspicion. Her judgment,

however, is a result of passive and largely unreflective habits that might not withstand scrutiny on careful reflection. Thus, Doris's emotion/judgment conflict is a rational conflict that falls short of extraordinary irrationality. In terms of her *attitudes* and their active and passive qualities, the conflict is less direct than a conflict between two simultaneously and actively cultivated or consciously maintained, but inconsistent attitudes. It has adequate distance from the limiting case on the irrationality spectrum to show that it falls within the scope of ordinary irrationality.

iv. A passive judgment/passive emotion conflict

In some emotion/judgment conflicts, we aren't engaged in much by way of the activities of reflection and control with respect to either the emotion or the judgment involved. For example, Nell is in the habit of judging that her research is good, though she has never been one to reflect on this judgment or subject it to scrutiny. This judgment is normally accompanied by pride. At the moment however, she is presenting her work at a conference and she has been made to feel shame about her work. Her commentator has pointed out an understandable, but potentially damaging, oversight that she's made. Normally, her sense of pride in her work wouldn't be shaken by such an exchange. However, in this instance, her commentator bears a slight but compelling resemblance in tone and manner to a family member whose strident criticisms of her poorer academic performances very often made her feel shame in childhood. She is surprised at this unusual sense of shame that she's experiencing, and she still judges that her research is something to be proud of, as it is her habit to do so.

Nell's conflict might involve more or less directness in the conflict between her attitudes' *contents* depending on whether she is experiencing a sense of shame that targets her present situation and current research project, or if she is experiencing a sense of shame that targets her work more generally. The former could be consistent with her judgment that her work is generally good if she can regard this instance as an exception to that. Whether her emotion/judgment conflict is more or less direct with respect to her attitudes' *contents*, it isn't direct with respect to the *attitudes* themselves. Neither her emotion nor her judgment bears significant active characteristics that could make the clash between her attitudes themselves extraordinarily irrational, as they would if they were both reflectively endorsed. Because her case lacks this aspect of directness, it falls short of extraordinary irrationality.

v. Degrees of control and the normative pressure to restore consistency

If there is any hesitation regarding the idea that emotion/judgments conflicts are irrational attitude conflicts, perhaps it's because there may be instances in which the conflict is due to a recalcitrant attitude over which we have little or no control or influence. There is a relationship between irrationality, normative pressure to conform to rational rules or standards, and the assignment of blame for failing to do so. Yet blame seems misplaced when it comes to recalcitrant states.

Here, it's important to acknowledge that our attributions of rationality and our assignations of blame do occasionally come apart. There are some conflicts that we may be less responsible for than others, even though they are irrational—perhaps even severely so. Degree of irrationality corresponds to features of the conflict, as described

above. Degree of responsibility corresponds to degree of control. As Nolfi points out, the wedge between irrationality and responsibility is borne out in both our evaluative practices and our interpersonal reasoning practices. If a person has a thoroughly and completely recalcitrant attitude, we excuse her from blame and we lower or erase the expectations that normally come with the espousal or demonstration of an attitude. Normally, we expect a query after reasons to be met with an attempt to provide them, and we expect attempts at persuasion to be given uptake. If a person's attitude is recalcitrant, we don't generally hold him to account for it.[178] Attitude conflicts are sometimes due to the presence of one or more recalcitrant attitudes. If there are recalcitrant attitudes involved in a conflict, one may have little or no power to resolve it. In that case, there can be little or no appropriate blame for one's failing to rationally resolve it, even though it is irrational.

Take Oliver's case for example. Oliver may have more or less control over his passive fear of turbulent flight, which is in conflict with his judgment that it is safe to fly and that turbulence is normal. If he is able to reflect on his situation and influence his attitudes so that they are consistent and equally supported by his reasons, there is normative pressure for him to do so. Say Oliver's fear is a result of a general sort of disposition to experience fear in response to a certain class of uncontrolled and unexpected motions that one experiences when one is driving in poor conditions and loses control of the vehicle. His fear response to the turbulence, then, is inappropriate because the motions in question aren't an indication that the plane is out of control. Oliver might be able to reflect on his situation and come to see it differently.

[178] Nolfi, "Which Mental States are Rationally Evaluable," 52-53.

Turbulence is to be expected on plane rides and it generally isn't correlated with significant danger. Additionally, Oliver might reflect on the fact that no one else on the plane seems worried, and the fact that the plane doesn't seem to be losing altitude. These reflections may move him to no longer be afraid. However, if no amount of reflection could resolve the attitude conflict because Oliver's fear doesn't respond to reasons, then the normative pressure for him to resolve the conflict is lessened. Generally, responsibility for an attitude requires having some sort of proprietary relationship over it involving influence or control. Thus, we tend to hold a person to be less responsible for attitudes over which little or no control is possible, though we hold their attitudes to be irrational.

Our judgments, too, can be more or less responsive to reflection. Take Doris's case for example. Doris has a tendency to automatically judge people who fit certain social stereotypes to be generally trustworthy, even though those stereotypes have little to do with any given individual's character or his intentions toward herself. As result, she judges that Floyd is trustworthy, even though she feels reluctant to do business with him for good reasons and on the basis of well-honed instincts. Doris bears responsibility for resolving the conflict with her emotions of fear and distrust and removing the irrationality insofar as she is capable of altering her habits of mind by reflecting on her reasons. Adjusting her errant tendencies will likely be more difficult if they are connected to deeply held beliefs bearing lots of connections to other attitudes. They will likely be easier to adjust if they are simply unreflective habits with no deep connections to keep them firmly rooted.

Though it may be appropriate to *let someone off the hook*, so to speak, for their irrational attitudes by acknowledging that responsibility is linked with control, a certain amount of caution is necessary. As I argued in chapter 3, there are some emotions that, on cursory examination, appear to be recalcitrant that aren't truly recalcitrant. It can be tempting to view an emotion as recalcitrant simply because it conflicts with a contrary judgment. However, emotions that conflict with judgments aren't always immovable, though they may not be moved by the same reasons that inform the deliverances of our judgment faculty. We are capable nevertheless of moving them with reflection. Adjudicating emotion/judgment conflicts often requires examining the reasons that support each attitude and using an expanded set of tools for reflection. Attention redirection, information elaboration, and imagination are often more effective influences of our emotional sensibilities than deliberation.[179]

4. Conclusion: Resolving Helm's dilemma

The examples discussed in section 3 show that emotion/judgment conflicts generally fall within the range of ordinary rationality, and a number of factors affect exactly where on the irrationality scale they fall.

[179] McHugh suggests that blame for emotions it limited specifically to the cognitive components of emotions. Affective components are sometimes exempted because they are non-cognitive, and thus beyond our control. See McHugh, "Epistemic Responsibility and Doxastic Agency," 151. I disagree. Emotions have multiple components including appraisals, action, tendencies, bodily change, cognitive styles, and feelings. These components are dynamic and they are all affected, in due course, by reflection. Reflection interferes with, adjusts, and potentially overrides the activation of an emotion. See chapters 2 and 3 of this dissertation.

The account I have given resolves Helm's dilemma by showing that emotion/judgment conflicts are attitude conflicts, and this need not imply that emotions/judgment conflicts are an extraordinary form of irrationality. All that's needed to resolve the dilemma is to show that, generally speaking, emotion/judgment conflicts aren't inconceivably irrational. Most, if not all, emotion/judgment conflicts will be variations on one of the three broad possibilities profiled in section 3, and will thus fit within acceptable bounds for irrationality that is perfectly ordinary. If the limiting case—a direct active judgment/active emotion conflict—is possible for creatures like us, it is highly unusual and unlikely to occur.

My approach to resolving Helm's dilemma is messier than the alternatives. Helm's and Brady's approach is comparatively simpler, neater, and more unified. They each purport to show that no emotion/judgment conflict could fail to be irrational, and that no emotion/judgment conflict could be incomprehensibly irrational by showing that all emotion/judgment conflicts are irrational in just the same way. Yet in order to provide these apparently simple accounts, both rely on false premises about the character of emotions and their relationship to our judgments. Helm relies on the notion that emotions are universally and broadly passive in nature, while judgments are active. This, I argued, isn't true in such a general fashion that it can support a complete account of emotion/judgment conflicts. Both emotions and judgments admit of active and passive features and varieties. Brady relies on the notion that judgments have rational authority that can never be challenged by a contrary emotion. This, too, is false. Our judgments are a sometimes in error, and it can be both rational and appropriate to resolve an emotion/judgment conflict in favor of an emotion.

Chapter 6

Emotion/Judgment Conflicts and Charitable Interpretation

When we're interpreting another person's attitudes, there are costs associated with interpreting the person's attitudes as irrational. For one, we sacrifice some of the plausibility of our interpretation. If we have a choice between an interpretation on which the person's attitudes are rational and one on which the person's attitudes are irrational, the interpretation on which the person's attitudes are rational tends to be more plausible. Accordingly, Quine observes, "one's interlocutor's silliness, beyond a certain point, is less likely than bad translation."[180] Secondly, when we allow irrationality into our interpretations of a person's attitudes, we sacrifice some of the ground on which any attitude ascription can stand. If we cannot assume that a person is, by and large, a rational being whose attitudes are manifest in her speech and behavior, we may not be able to interpret her speech and behavior in terms of her attitudes at all. Thus, interpretive success depends, to a large degree, on our upholding an assumption of a person's general rational consistency. For either or both of these reasons, most philosophers find some version of the principle of charity to be a reasonable constraint on attitude interpretation. Roughly stated, the principle is: On balance, prefer an interpretation of a person's attitudes which minimizes irrational inconsistency. Though occasional attitude conflicts are of course possible for ordinary human agents, we ought to take care that we don't attribute irrationality to a person if there is a better, more rational alternative available.

[180] W. V. O. Quine, *World and Object* (Cambridge, MA: MIT Press, 1960) 59.

The principle of charity is frequently invoked in the context of discussions about how to attribute attitudes to one who is experiencing an emotion/judgment conflict. Emotion/judgment conflicts occur when our emotions and judgments are at odds with each other: One experiences jealousy, though at the same time, one judges that there is no threat to one's relationship, or one experiences shame over something one judges worthy of one's pride. We normally expect a person's emotional reactions to fall in line with the other speech and behavior that we would usually take to be indicative of her beliefs, judgments, and intentions. In emotion/judgment conflict situations, however, they don't. When experiencing an emotion/judgment conflict, a person will exhibit speech and behavior that fails to fit a unified and consistent attitude-profile. For example, someone who experiences jealousy though they judge that there is no threat might be angry at their partner's social behaviors for reasons that they apparently disavow. This presents an interpretive challenge. The principle of charity enjoins us to avoid irrational attitude ascriptions—particularly those involving incoherent attitude sets—in order to avoid the costs associated with too hasty ascriptions of irrationality. But if emotions are treated as attitudes, then it looks like there is no way to attribute rational attitudes in these cases. Hence, the principle of charity is taken to show that emotions are not attitudes.

In this chapter, I argue that the principle of charity does not require what the argument from charity suggests that it requires when we are interpreting the attitudes of one who is exhibiting signs of emotion/judgment conflict. The principle of charity constrains attributions of irrationality in order to preserve an assumption of rationality, and this in turn helps us to avoid the costs associated with irrational attitude ascriptions.

However, in some cases, ascriptions of irrationality are appropriate, the principle of charity notwithstanding.

I begin with a discussion of the most prominent version of the argument from charity: Patricia Greenspan's argument in "Emotions as Evaluations."[181] Her widely cited argument will serve as proxy for a larger trend in the literature for the purposes of this paper. In section 1 I offer a brief characterization of the principle of charity and the arguments that support it. I enumerate several constraints it places on attributions of irrationality. In section 2 I discuss the principle of charity's implications for emotion/judgment conflicts. I demonstrate several failings of Greenspan's interpretation of emotion/judgment conflicts. In section 3, I argue that the principle of charity does not preclude attributions of irrationality in cases of emotion/judgment conflict across the board. I then show that my characterization of the nature of emotion/judgment conflicts does not violate any of the constraints I lay out in section 1.

1. The argument from charity

The most often cited version of the argument from charity comes from Patricia Greenspan's article "Emotions as Evaluations."[182] She illustrates her argument with the following example involving a person who is afraid of dogs. For ease of discussion, I give the protagonists names: Agnes and Millie. Agnes had a very bad experience with a dog in the past, and now she can't help but be afraid of any dogs she encounters. One day, she comes across Millie, a friendly canine senior citizen with declining health and

[181] Greenspan, "Emotions as Evaluations," 158-169.
[182]Ibid. The argument is reprised in Greenspan, *Emotions and Reasons,* 18.

very few teeth. She judges that Millie is no significant threat to her. She also judges that it would be safe for her to approach Millie, and safe for others to approach Millie too. She wouldn't warn others not to approach if they were to move in the dog's direction. Nevertheless, she shrinks away from Millie due to her fear, and she would go to great lengths to avoid Millie if the necessity of approaching her could possibly be circumvented.[183] If Agnes' fear of Millie represents the attitude that Millie is dangerous and that it's not okay for her to approach Millie, born out in her commitment to avoidance stemming from her fear, Agnes is experiencing some significant inter-attitudinal conflict: She's committed, by her judgment, to Millie's harmlessness at the same time that she's committed, by her emotion, to Millie's danger. She is committed, by her emotion, to avoiding Millie, but due to her judgment, she is committed to refraining from warning others to stay away. Thus interpreted, fearing Millie while judging that she is not a threat is patently irrational. It represents a set of contradictory attitudes.[184]

Greenspan comments, "Logical incoherency is possible, of course; but I am assuming that the agent is functioning quite rationally in general."[185] Emotion/judgment conflicts are not unusual. Nor do they generally interfere with our ability to recognize what our rational commitments are given our judgments, though these commitments conflict with our emotions. Thus, she maintains that the principle of charity ought to apply. Instead of attributing an irrational set of attitudes to a person who has an emotion/judgment conflict, a better interpretative strategy would yield the verdict that

[183] Greenspan, "Emotions as Evaluations," 162.
[184] Ibid., 162-3.
[185] Greenspan, *Emotions and Reasons,* 18.

she has rational attitudes. Greenspan's interpretive strategy preserves logical coherency in our interpretation in such cases by treating judgments as attitudes, but emotions as non-commitment oriented mental states. Hence, the principle of charity requires that we treat emotions as non-attitudes.

Greenspan's argument has been widely influential in the philosophy of emotion. D'Arms and Jacobson endorse her argument, observing that if we view emotions as intentional attitudes, we'll be forced to attribute "peculiarly conflicted beliefs" to those who are experiencing emotion/judgment conflicts.[186] Tappolet similarly rejects attributions of contradictory judgments to a person experiencing emotion/judgment conflicts. She writes:

> …if one assumes that emotions involve an evaluative judgement, one would have to attribute inconsistent or even contradictory judgements to the person who experiences the emotion. For instance, she would judge that the object of her fear is fearsome, while also judging that it is not. But whatever irrationality is involved in recalcitrance, it seems to be of a less acute species than what is involved in inconsistent or contradictory judgements.[187]

Michael Brady, too, accepts Greenspan's reasoning and claims that her argument is a significant challenge to views on which emotions involve an intentional attitude such as a judgment. He charges: "Since the only reason the judgementalist seems to provide for this attribution [of irrationality in cases of emotion/judgment conflict] stems from their adherence to the judgementalist theory, their explanation of recalcitrant emotions is undermined."[188]

---

[186] Justin D'Arms and Daniel Jacobson, "The Significance of Recalcitrant Emotion," in *Philosophy and the Emotions*, edited by Anthony Hatzimoysis (Cambridge, UK: Cambridge University Press, 2003) 129-130.
[187] Tappolet, "Emotions, Perceptions, and Emotional Illusions," forthcoming.
[188] Brady, "The Irrationality of Recalcitrant Emotions," 415.

Each of the above authors concludes that emotions do not belong in the class of intentional attitudes along with beliefs, judgments, desires, and intentions. They argue that if we can avoid interpreting emotion/judgment conflicts as irrational inner conflicts, we ought to. We can avoid it, they claim, by assimilating emotions to a class of non-commitment-oriented mental states that includes thoughts, perceptions, and inclinations.[189] These are mental states that do not necessarily involve reason-based commitments, as our beliefs, judgments, and intentions do. They are not, therefore, attitudes, and as nonattitudes they cannot create inter-attitudinal incoherence.

Surprisingly, neither Greenspan nor those who endorse her argument provide much further discussion of the principle of charity, though it serves as a foundational assumption for the argument that emotions are not attitudes. The principle of charity is not, however, uncontroversial. Is it also, arguably, not obvious that it applies to cases of emotion/judgment conflict in the manner that Greenspan suggests. In the next section, I offer a discussion of the possibility of irrational attitude conflicts and the circumstances in which it might be appropriate to countenance such conflicts.

2. The principle of charity and the possibility of human irrationality

I mentioned two common rationales for accepting the principle of charity at the beginning of the paper. Here, I develop those rationales and the constraints they place on attributions of irrationality. I show that the principle of charity does not require us to

[189] Greenspan suggests thoughts, Tappolet suggests perceptions, and Brady suggests inclinations. See Greenspan, *Emotions and Reasons*, 158-169; Tappolet, "Emotions, Perceptions, and Emotional Illusions," forthcoming; and Brady, "The Irrationality of Recalcitrant Emotions," 413-430.

force an interpretation of rationally consistent attitudes in every case of emotion/judgment conflict. Briefly, those two rationales were: (1) Interpretations that include irrational attitude attributions are simply less plausible than ones on which the person interpreted is rationally consistent, and (2) Maintaining an assumption that a person is, by and large, rationally consistent is arguably necessary for attitude interpretation to be possible. If this assumption is undermined, we may not be able to attribute attitudes to the person at all. The principle of charity preserves the presumption of rationality, and so it is a reasonable constraint on attitude interpretation.

The latter of these two is associated with interpretationism. Interpretationism is a common methodological approach in the philosophy of mind, the premise of which is that we can learn something about a person's mental states by interpreting her speech and behavior in light of the beliefs and other attitudes that would make her speech and behavior intelligible. We can safely assume, for example, that if Lucy announces "I'm going to read Dickens today" before walking to a nearby bookshelf and taking *Little Dorrit* off the shelf, she believes that *Little Dorrit* is a Dickens novel, she intended to take it from the shelf, and that she desires to read it. The basic tenets of interpretationism are: (1) Our attitudes are reason-based;[190] and (2) Our attitudes fit together into a largely coherent, rational whole.[191] Taken together, these claims suggest

[190] Daniel C. Dennett, *The Intentional Stance* (Cambridge, MA: MIT Press, 1987) 21-22; Donald Davidson, *Problems of Rationality* (Oxford: Oxford University Press, 2004) 164.

[191] Dennett claims that if we can assume that a person has an attitude, we can also assume that they have an attitude toward whatever the attitudes implications are (within the limits of what's pertinent given the person's current situation). See Dennett, *The Intentional Stance*, 20-21. Davidson claims that to have one attitude, you have to have many others that support it and confer meaning on it. See Donald Davidson, "Rational

that we are equipped to interpret each other's attitudes by means of observing each other interact with our shared environment. We can infer a person's reasons from her speech and behavior, and hence, her attitudes (or vice versa).[192]

These two rationales for adopting the principle of charity support somewhat different constraints on irrational attitude attribution. The plausibility argument provides comparatively weak and nebulous constraints on irrational attitude attributions. By contrast, when bolstered with interpretationist commitments, the principle of charity provides us with several specific and stringent guidelines delimiting when it could be permissible to attribute irrationality.

The plausibility argument places fewer constraints on attitude ascriptions than the interpretationist argument because it simply tells us that rational attitude ascriptions tend to be more plausible than alternative interpretations on which a person's attitudes are irrational. We expect human reasoners to have more or less rationally consistent beliefs. Cases in which a person fails to be rationally consistent are, without extenuating circumstances to explain them, surprising. As a rationale for the principle of charity, the plausibility argument supports a general preference for rational interpretations, but it leaves room for this general preference to be overridden by any number of intervening factors. If there are good reasons to think that a person is likely to make errors in a particular type of circumstance (for example when under duress; when unconscious motives are in play; when heuristics and biases interfere with logical reasoning; etc.), it

---

Animals," *Dialectica* 36.4 (1982): 321; and Donald Davidson, "Radical Interpretation," *Dialectica* 27.1 (1973): 324.

[192] For a helpful characterization of interpretationism, see William Child, *Causality, Interpretation, and the Mind* (Oxford: Oxford University Press, 1994) 8-13.

may be reasonable for us to attribute errors to the person when those circumstances obtain. We should, perhaps, take precautions even in these circumstances not to attribute irrationality to a person too hastily, but if careful reflection on the evidence supports an attribution of irrationality, it's not obviously unreasonable to make such attributions.

The interpretationist argument for the principle of charity provides us with greater constraints on attitude ascription than the plausibility argument alone. For an interpretationist, the principle of charity constrains the activities of translation and attitude interpretation because it's a necessary assumption for interpretation to be possible. This is most apparent when we consider the activity of radical interpretation. A radical interpreter attempts to interpret the attitudes of another person without the help of a prior theory about the person's background attitudes and commitments. In order to successfully interpret another person under these conditions, the interpreter must assume, as a starting point, that the person she interprets has attitudes that follow logical rules and are largely true. This is a starting point from which she can develop a first approximation. Then, the interpreter can compare her theory to the person's speech and behavior and recalibrate her theory to better fit with it, continuing to make adjustments as ongoing speech and behavior are observed.[193]

Unsurprisingly, it often proves very difficult to produce an interpretation encompassing all of a person's attitudes that eliminates all inconsistency and error. Both common sense and empirical evidence strongly suggest that we do sometimes make mistakes. The principle of charity requires that we treat the possibility of mistakes very

[193] See J. E. Malpus, "The Nature of Interpretive Charity," *Dialectica* 40.1 (1988): 19.

carefully. Inconsistencies and errors are obstacles for interpretation because they don't lend themselves to the construction of an interpretation that posits a consistent and fully integrated network of attitudes. Thus, inconsistency and error attributions are to be avoided whenever possible. When it is not possible—especially if the inconsistencies and errors are surprising or egregious in nature—we owe special stories explaining how such inconsistencies could arise.[194]

Additionally, any attributions of irrationality that are made must meet constraints that preserve the fundamental assumptions of interpretationism. The specific constraints that interpretationism places on attributions of irrational attitudes are three in number. First, the number of irrational attitudes a person could have is subject to limits. Second, a person's irrational attitudes must not be wholly unconnected to his or her other attitudes or commitments. Third, irrational attitudes must not be produced by systematic reasoning failures. I elaborate on these three constraints on attributions of irrational attitudes below. In section 3, I argue that none of these constraints, nor any associated with the plausibility argument, preclude attributions of inconsistency in cases of emotion/judgment conflict.

The first constraint on attributions of irrationality is that the number of such attributions cannot be too many. If a theory of interpretation attributes a great number of errors to a person, it compromises our ability to understand the person's attitudes. At some point, too many errors and inconsistencies threatens our ability to interpret the person's speech and behaviors as expressions of rational attitudes at all. This is because

[194] See Dennett, *The Intentional Stance*, 20. Here, Dennett explains the need for special stories to explain irrationality.

the presence of too many errors and inconsistencies could represent a breakdown amongst the sorts of logical connections that would make attitude attributions supportable.[195]

The second constraint is that the attitudes involved cannot be unconnected to other attitudes that make it meaningful. This constraint stems from the fact that attitudes lack meaning when they lack connections to other attitudes. Davidson puts the point this way:

> We cannot intelligibly attribute the thought that a piece of ice is melting to someone who does not have many true beliefs about the nature of ice, its physical properties connected with water, cold, solidity, and so forth. The one attribution rests on the supposition of many more—endlessly more.[196]

When we're interpreting a person's attitudes, it would be deeply problematic to attribute an attitude to him if it is not supported by a network of attitudes that make it intelligible. If the attitude bears no apparent relationship to his other attitudes, it shouldn't be attributed to him. Malpus provides a helpful illustration of this claim in his discussion of the principle of charity. He writes:

> Suppose we interpret an utterance of a speaker as expressing the belief that 'ice melts when left out in the warm sun'. However, in further discussion the speaker says things which we take to express beliefs to the effect that ice does not melt when placed in a hot oven and that ice is produced through the compaction of sugar crystals. While the belief that ice melts when left in the warm sun is, generally speaking, true, and the latter two beliefs are generally false, when these beliefs are held by the same speaker at the same time then it becomes unclear just what is being believed. For instance, is it ice which these beliefs are about?[197]

---

[195] Davidson, *Problems of Rationality*, 184.
[196] Ibid., 183. See also Donald Davidson, "Incoherence and Irrationality," *Dialectica* 39.4 (1985): 346.
[197] Malpus, "The Nature of Interpretive Charity," 20.

In a case such as this one, whatever appearances suggested that the speaker believes that 'ice melts when left out in the warm sun,' the totality of the evidence recommends against attributing the belief to the speaker. Without the appropriate structure of supporting attitudes, belief attribution would be tenuous at best.

The third constraint on attributions of irrationality—and the most contentious one—is that inconsistencies and errors cannot be the result of systematic reasoning failures. Defenders of this constraint argue that the vast majority of inconsistencies and errors are mistakes due to inattention, confusion, distorted perception, or failure to follow correct procedure.[198] That is, we can presume generally that mistakes are failures to reason in accordance with our ability. When our mistake becomes apparent, we are capable of revising our mistaken attitude to accord with sound rational principles. If this view is correct, it would be a mistake to attribute forms of irrationality that represent systematic problems with our rational faculties.

This constraint is highly contentious, however, because there is a growing body of literature in psychology arguing that normal human beings do exhibit certain kinds of systematic irrationality. The results of Wason selection task experiments are illustrative. Wason selection tasks test individuals on their intuitive understanding of the logic of conditionals. The selection tasks vary, but they follow a pattern similar to the following: Subjects are presented with four cards, each with a letter on one side and a number on the other. They cards are placed in the following sequence: O, C, 4, 5. They are asked to determine which of these cards they must turn over to test the truth of the claim, 'If

[198] Dennett, *The Intentional Stance*, 87. See also L. Johnathan Cohen, "Can Human Irrationality be Experimentally Demonstrated?" Behavioral and Brain Sciences 4.3 (1981): 317-370.

there is a vowel on one side, there is an even number on the reverse.' Very few participants in these studies answer correctly. Instead of the O-card and the 5-card, they select only the O-card or the O-card and the 4-card.[199]

To resist the notion that these results are indicative of a mistake, proponents of the constraint point to evidence that we do have the capability to perform the Wason selection task correctly. There are variations of the Wason selection task that participants in the studies usually answer correctly. For example, if the conditional to test is 'If a person is under 21 years of age, he or she is not drinking alcohol' most participants will choose the correct cards to confirm or disconfirm the rule.[200] In light of these results Cohen comments, "subjects who reason fallaciously [on the Wason selection task] …need not be supposed to lack the correct deductive 'program.' The subjects merely fail to recognize the similarity of their task to those familiar issues in which they have profited by using the deductive procedure."[201]

However, critics of the systematic reasoning failure constraint offer reasons to doubt an interpretation of the Wason selection task results such as Cohen's. In some iterations of the task, participants who answered incorrectly were instructed to uncover the correct cards and brought to realize that the rule was in fact false. Participants would often persist in thinking that their wrong answer was correct even after the correct answer was supplied to them and explained. This suggests that they were not simply

---

[199] Stephen Stich, "Could Man be an Irrational Animal," *Synthese* 64.1 (1985):116-118; Edward Stein, *Without Good Reason: The Rationality Debate in Philosophy and Cognitive Science* (Oxford: Oxford University Press, 1996) 90-92.
[200] See Stein, *Without Good Reason*, 88 for interpretations of the ability exhibited cases like this one.
[201] Cohen, "Can Human Irrationality be Experimentally Demonstrated?" 324.

failing to grasp the need to apply a particular principle because they've misinterpreted the task and applied the wrong principle. Showing participants the correct principle fails to lead them to a correct understanding of the problem or to influence their judgments.[202] For this reason, some proponents of the principle of charity advocate a more modest version which allows that we can make attributions of irrationality that are in line with empirical evidence about ordinary human reasoning failures.[203]

Though there is some compelling evidence that human beings are prone to certain kinds of systematic rational failures, it may be prudent to assume that it may yet be too soon to be sure. To quote Edward Stein's exceedingly conservative conclusion to his book *Without Good Reason: The Rationality Debate in Philosophy and Cognitive Science*,

> To determine what is required to be rational, we need to know how various factors constrain our reasoning. Even if we know what is required for humans to be rational, to determine whether humans are in fact rational, we need to know more than we do about, for example, human reasoning behavior, human neurophysiology, and human evolutionary history. Anyone who has confidently asserted either that humans are rational or that humans are irrational does so on the basis of incomplete empirical evidence or unsupported conceptual claims; in other words, she has taken a strong stand on the question of human irrationality without good reasoning.[204]

Thus, to be prudent, I will provisionally allow that the principle of charity precludes attributions of irrationality that represent systematic reasoning failures, though I acknowledge that it could turn out to be a false requirement.

---

[202] Stein, *Without Good Reason*, 92.
[203] See David K. Henderson, "An Empirical Basis for Charity in Interpretation," *Erkenntnis* 32.1 (1990): 83-103. See also Paul Thagard and Richard E. Nisbett, "Rationality and Charity," *Philosophy of Science* 50. 2 (1983): 250-267.
[204] Stein, *Without Good Reason*, 277.

In the next section, I return to the question of emotion/judgment conflict. I have provided a discussion of the principle of charity and the possibility of irrationality in order to show that, at the very least, it is not obviously the case that the principle of charity precludes attributing irrationality to one who is experiencing emotion/judgment conflict. On the contrary, I will argue that attributions of irrationality may be called for in cases of emotion/judgment conflict. I will show that attributions of irrationality can be made in cases of emotion/judgment conflict without violating any of the constraints outlined above.

3. The principle of charity's implications for emotion/judgment conflicts

Greenspan argues that in cases of emotion/judgment conflicts, we should not attribute conflicting intentional attitudes to a person. She supports her argument with an appeal to the principle of charity. The principle of charity, as Greenspan characterizes it, requires a "presumption of basic rationality" that we employ "in attempting to explain the various phenomena (experiential and behavioral) that we are actually confronted with" when we're interpreting one another's attitudes.[205] Our emotional reactions, which are usually in line with our judgment, are usually good indicators of our attitudes. However, in emotion/judgment conflicts, she claims they are not. She suggests that, due to the traumatic nature of Agnes' experience, Agnes may have formed some false beliefs or judgments about the danger posed to her by dogs, consciously or not. These past beliefs may have produced her current overgeneralized emotional reactions that lead her to fear dogs whether she would judge them to be dangerous or not. However,

[205] Greenspan, "Emotions as Evaluation," 163.

in the present case in which she acknowledges her emotion to be wholly groundless, we shouldn't attribute false or inappropriate attitudes to her. She writes,

> Now that [Agnes is] functioning rationally in most cognitive respects, it seems more plausible to trust my view that {Agnes'} present emotional reactions are groundless. The causal influence of *past* beliefs, via present associative links, would seem to be enough to explain them—and to drive a small wedge between emotion and judgment.[206]

On Greenspan's view, Agnes' emotion is an indication of a past attitude, but at present, it is just a persistent thought—"Millie is going to bite me!"—due to her emotional habits.[207] It bears no logical or practical commitments.

Greenspan supports her interpretation of the case by noting that, ordinarily, a person who is experiencing an emotion/judgment conflict is "functioning quite rationally in general."[208] That is, emotion/judgment conflicts don't generally compromise our ability to exercise rational judgment and have attitudes that are by and large consistent with one's judgment. Greenspan's interpretation of Agnes' case "rests on the assumption that even a highly disturbing object-centered evaluation need not impair one's judgment."[209] Agnes fears Millie, but judges that she poses no threat. She is also committed to the sorts of attitudes that follow from her judgment, i.e. that no one else is in danger in Millie's presence. Agnes' fear is not grounded in the rational commitments that she takes herself to have. Her fear is thus *not her attitude* on Greenspan's view, though it is a mental state—an evaluative thought—that is irrational

[206] Ibid.
[207] Ibid.
[208] Greenspan, *Emotions and Reasons*, 18.
[209] Greenspan, "Emotions as Evaluations," 163.

because it isn't appropriately grounded in her rational commitments. Agnes doesn't hold the sorts of attitudes which would make such an evaluative thought appropriate.[210]

In this section, I will challenge two key points in Greenspan's argument: the notion that a person who is experiencing an emotion/judgment conflict is functioning rationally in general, and the notion that we can make attitude attributions that preserve rationality in cases of emotion/judgment conflict by holding that emotions are not judgments or any other sort of intentional attitudes. I will argue that correct reasoning does not guarantee attitude coherence and, hence, rationality. Thus, Greenspan's argument fails to demonstrate that a person experiencing emotion/judgment conflict is functioning rationally. I will also argue that the claim that emotions are not our attitudes introduces further problems for rational interpretation. It introduces the problem of non-agential forces that drive some portion of our speech and behavior. From the interpretationist point of view, Greenspan's solution to the problem of emotion/judgment conflict puts us in a very similar position to the one it was supposed to avoid. That is, if we are steered by forces that are non-agential with some regularity, it would seriously threaten the fundamental assumption behind interpretationism—the notion that we can gain knowledge of each other's attitudes by carefully observing each other's speech and behavior because our speech and behavior are appropriately connected to our attitudes.

Contrary to Greenspan's view, I will argue that emotion/judgment conflicts represent attitude conflicts. I will take for granted that the principle of charity reasonably constrains attributions of irrationality in the ways discussed in the previous

[210] Ibid., 164-165.

section. But, I will argue, none of these constraints precludes attributing irrationality to one experiencing an emotion/judgment conflict.

When Greenspan posits that a person experiencing emotion/judgment conflict is functioning rationally in general, she is making reference to the fact that the person is reasoning correctly and making plans that accord with her judgment. When Agnes is making plans, she might anticipate being emotionally disturbed by Millie's presence and thus try to avoid any encounters with her, as is prudent. Nevertheless, she will, in large part, premise her plans on the notion that Millie is no real threat to herself and others.[211] Though she feels conflicted, she recognizes what follows from her judgment, and hence, what her rational commitments are. Emotion/judgment conflicts do not typically undermine our ability to reason in accordance with rational rules or to make plans that accord with our judgments.

The assumption that reasoning in this manner makes a person rational is, I contend, a false assumption for reasons I will enumerate presently. A preliminary observation to make is that, from an interpretationist point of view, the assumption of rationality that the principle of charity demands is supposed to preserve interpretability and predictability. It's not clear that its application in emotion/judgment conflict cases serves this purpose. In emotion/judgment conflicts, our speech and behavior can pose some difficulties for the purposes of interpretation and prediction. Since speech and behavior of one who is experiencing emotion/judgment conflicts is likely to be disjointed and inconsistent, it will be difficult to discern from the person's total set of speech and behavior precisely what her rational commitments are, and hence, what her

[211] Ibid., 164.

attitudes are. We would normally expect that one who judges that Millie, the aged, toothless dog, is nonthreatening would speak and behave in a manner that accords with that judgment. She would not worry about Millie or go out of her way to avoid her. Since Agnes' emotions conflict with her judgment, however, we can expect her to exhibit some speech and behavior that do not conform to her judgment. At the same time that she asserts that Millie is no danger, Agnes' fear might cause her tremble, wince, cry out, or freeze. In addition to these subtle behaviors, she might take more overt actions. She might avoid the room that Millie is in even if it is a great inconvenience to herself and others for her to do so. If she stays in the room, she may carefully observe every move of Millie's because she cannot abandon the worry that Millie might become aggressive or harmful. She might, if asked, even describe Millie's manner as suggestive of a possible underlying aggressive tendency because her fear primes her to interpret any movement or gesture of Millie's that might be so-construed as threatening. She might also experience a momentary sense of panic if Millie were approach her, and then betray her surprise when Millie gives her a friendly nuzzle rather than a warning bark. Thus, her speech and behavior generally suggest inner conflict even though her reasoning faculties are functioning correctly.[212] Greenspan's interpretation of Agnes' attitudes therefore fails to make sense of her speech and behavior because it defies our expectations of one whose attitudes are in line with the notion that Millie is simply not a threat.

---

[212] For more on emotions' influence on our attitudes and behaviors, see Philippot et al., "Cognitive Regulation of Emotion," 71-97; Brady, "Emotion, Attention, and the Nature of Value," 52-71; and Antoine Bechara, "A Neural View of the Regulation of Complex Cognitive Functions by Emotion," in *The Regulation of Emotion*, edited by Pierre Philippot and Robert S. Feldman, 3-32, Mahwah, NJ: Lawrence Erlbaum, 2004.

There are good reasons to doubt that correct reasoning alone issues in consistent attitudes generally. We can, for example, reason to form an all-things-considered judgment that we ought to ϕ, but fail to intend to ϕ. This need not indicate any failure of our ability to reason according to rational rules. Our judgment is simply rooted in reasons that are different from the ones that support our intentions. Davidson explains such attitude conflicts in terms of mental partitioning. Conflicts between partitioned parts of the mind occur, he suggests, when one attitude in the conflict has a cause that is not a proper reason for the attitude. A desire in one partitioned part of the mind, for instance, might cause an errant intention in another partitioned part of the mind in cases of weakness of will or an errant belief in cases of self-deception.[213] The desire motivates efforts to support the errant belief or intention with reasons that are partitioned from the ones that support the judgment.

While irrationality due to partitioning can go unnoticed, Davidson suggests that conflicts can persist within a person who suspects, or is fully aware, that he harbors an inner conflict of attitudes.[214] The partitioned deliberative resources need only to be "functionally isolated."[215] For example, a woman finds lots of illegal drug paraphernalia hidden in the sock drawer of her teenage son's room but does not want to be the kind of mother who would believe that her child would use illegal drugs. So, she intentionally avoids the sock drawer in her son's room so as not to confirm what her continued act of avoidance suggests she already knows. She's simply constructed a mental boundary that

[213] Davidson, *Problems of Rationality,* 179.
[214] Ibid., 179.
[215] John Heil, "Minds Divided," *Mind* 98.392 (1989): 575.

allows her to, at the same time, maintain the belief that her son would not use illegal drugs.

A person who harbors partitioned mental structures can be functioning rationally in the sense that her capacity to reason and plan according to the rules of logic is not impaired. Within each partitioned domain, her attitudes are arrived at by a well-functioning reasoning faculty. Yet, she is irrational because she harbors inconsistent attitudes.

Another reason to think that that our attitudes can conflict without necessarily compromising our ability to reason according to rational rules is that our reasoning faculties do not issue in all, or even the majority, of our attitudes. Many of our attitudes issue from type 1 processes. Recall that type 1 processes are the fast, high capacity, and low effort processes that we rely on for the bulk of our cognitive functions. The other type 2 processes are by comparison slow, low capacity, and high effort, and they are of necessity engaged sparingly. Explicit, deliberative reasoning activities are of the latter type.[216] Sometimes, multiple cognitive processes are engaged at once. Type 1 processes run parallel to each other and to type 2 controlled reasoning and deliberation processes. As a result, the outputs of these independent processes occasionally diverge.

At times, a type 2 reasoning process will be capable of suppressing a type 1 process from issuing its usual outputs. However, suppression isn't always possible. Type 2 interventions require time, cognitive resources, and motivation that we can't sustain all the time.[217] Consequently, the piecemeal type 2 overrides that we are capable

[216] See Chapter 5, footnote 163.
[217] Daniel Kahneman and Shane Frederick, "Representativeness Revisited: Attribute Substitution in Intuitive Judgment," in *Heuristics and Biases*, edited by Thomas

of supporting may not be sufficient to address the deeper conflict. Without efforts to retrain or implement an automatic regulating routine on type 1 processes so that they accord with our type 2 processes and their outputs, there will continue to be discrepancies between these processes as they are each engaged at alternate moments.[218]

For example, implicit association tests demonstrate that, due to a variety of social factors, persons who consistently reason in accordance with non-racist attitudes using type 2 processes may exhibit racist implicit attitudes that accord with commonly held stereotypical associations with particular races when using type 1 processes. Implicit association tests measure the strength of one's implicit associations by measuring facilitation and inhibition of our ability to match racialized names and faces to both positive and negative words. Even though one reasons in accordance with non-racist attitudes, when one does not engage in explicit reasoning, but relies on one's automatic associative faculties, one's racist attitudes will be manifest in one's speech and behavior. When one does engage in explicit reasoning, but not carefully, one's implicit attitudes may still influence the course of one's reasoning processes by acting as inputs to explicit judgment in some cases. Implicit biases are thus an impediment to

---

Gilovich, Dale Griffin, and Daniel Kahneman (Cambridge: Cambridge University Press, 2002) 57-8.

[218] For discussion of the implications of conflicts between type 1 and 2 processes, see Leland F. Saunders, "Reason and Intuition in the Moral Life: A Dual-Process Account of Moral Justification," in In Two Minds: Dual Processes and Beyond, edited by Jonathan St. B. T. Evans and Keith Frankish (Oxford: Oxford University Press, 2009) 339; and Jonathan St. B. T. Evans, "On the Resolution of Conflict in Dual Process Theories of Reasoning," *Thinking and Reasoning* 13.4 (2007): 321-339.

maintaining consistent attitudes about race that accord with one's non-racist judgments.[219]

Thus, a person who reasons according to rational rules when she puts her mind to it may sometimes nevertheless exhibit attitudes that are conflicted. We sustain conflicted attitudes when our type 2 reasoning processes produce one set of attitudes and commitments and our faster type 1 processes produce another conflicting set of attitudes and commitments. The assumption that a person is rational if there is no breakdown in her reasoning ability is thus false. Conflicts can arise due to discrepant attitude forming processes operating either in parallel or in alternating sequences.

An objector might challenge the notion that conflict between type 1 and type 2 process outputs are really conflicts of attitudes. It's tempting to assume that only our conscious, reasoned convictions count as our attitudes. This is a mistake. Very few of our attitudes are arrived at by an explicit reasoning process. A very large portion of our speech and behavior issues from our type 1 attitudes.[220] Thus, if something is an attitude only if it issues from a reasoning process, most of our speech and behavior aren't connected to our attitudes because type 1 processes are the source of most of our speech and behavior. To maintain the commonsense notion that human beings are the sorts of creatures that generally do things for reasons, this conclusion would be unacceptable.

---

[219] See Timothy D. Wilson, Samuel Lindsey, and Tonya Y. Schooler, "A Dual Model of Attitudes," *Psychological Review* 107.1 (2000): 102; Eliot R. Smith and Elizabeth C. Collins, "Dual-Process Models: A Social Psychological Perspective," in *In Two Minds: Dual Processes and Beyond*, edited by Jonathan St. B. T. Evans and Keith Frankish (Oxford: Oxford University Press) 197-212; and Thierry Devos, "Implicit Attitudes 101: Theoretical and Empirical Insights," in *Attitudes and Attitude Change*, edited by William D. Crano and Radmila Prislin (New York: Psychology Press, 2008) 62-68.

[220] See Chapter 5 footnotes 164 and 165.

We can support the view that type 1 thinking issues in attitudes by observing that type 1 outputs, though automatic, are shaped by experience and effort. Type 1 processes draw from our implicit memory stores and perform automatized functions, many of which come from the transfer of repeated explicit processes from working memory into intuitive ones that obviate the need for explicit reasoning procedures.[221] Therefore, type 1 outputs ought to be viewed as being rooted in our reasons. The reasons that they represent have been embedded in a structure that produces attitudes without an explicit reasoning process. Though those processes are unconscious and automatic, their outputs issue from our background attitudes and our mental character.

To deny that emotion/judgment conflicts are conflicts between attitudes is, moreover, to make the instability and potential for unpredictable speech and behavior due to emotion/judgment conflicts more puzzling than it needs to be. If one disavows an emotion, and it is therefore not one's attitude, then it is a non-attitudinal and hence non-agential influence on speech and behavior. Insofar as we take the purpose of avoiding an ascription of irrationality in cases of emotion/judgment conflicts to be, in part, to avoid undermining our ability to interpret a person as a successful agent, this is problematic. It suggests that a person's speech and behavior are steered by something other than her reasons and her attitudes.

Greenspan suggests that emotions are irrational when they conflict with judgments because they represent groundless evaluative thoughts. The irrationality that emotion/judgment conflicts represents, it seems, redounds to the emotion-relevant

[221] See Smith and Collins, "Dual-Process Models: A Social Psychological Perspective," 202.

evaluative thought itself rather than to the person, who Greenspan argues to be generally functioning rationally. If this is correct, there is also a sense in which the evaluative thought represents a non-agential force or motivation within the person's psychology. The emotion, we are told, is not rooted in the person's reasons or commitments, and yet it issues in some speech and behavior that accords with the groundless evaluative thought associated with the emotion in spite of the person's generally functioning rationally. Though thoughts can usually be present to mind without any significant causal impact, Greenspan acknowledges that these emotional ones do. To be led by non-agential forces such as this is to make the speech and behavior it results in largely uninterpretable in terms of agential intentional attitudes because it is speech and behavior that is not appropriately rooted in the person's reasons and commitments.[222] This is a significant problem from the interpretationist point of view. If non-agential speech and behavior is a widespread phenomenon, it's a threat to the entire edifice of interpretationism, which assumes that we are reason-guided and that our reasons are manifest in our speech and behavior.

If we are to preserve the connection between reasons, agents, and observable speech and behavior, we must acknowledge that emotions are a person's attitudes. Emotion/judgment conflicts are thus, I submit, best interpreted as attitude conflicts. If emotions are attitude conflicts, we can avoid problems stemming from the dissociation of a person and her reasons from her emotions and her emotion-related speech and behaviors that is called for by Greenspan's view. There are good reasons, furthermore,

[222] Davidson, *Problems of Rationality*, 175, 180.

to think that emotion/judgment conflicts are attitude conflicts rooted in competing commitments harbored within the same person. I turn to these next.

4. Emotion/judgment conflicts: a charitable interpretation

Emotion/judgment conflicts could result from either mental partitioning, or discrepant mental processing outputs. These discrepancies could be between type 1 and type 2 process outputs or between two type 1 processes' outputs. Most of our emotion/judgment conflicts are probably the latter. Agnes' emotion/judgment conflict could be explained by mental partitioning if Agnes harbors a desire that motivates her to remain afraid of dogs. Perhaps her father's dog attacked her as a child and she derives a vengeful pleasure from her father's remorse over her persistent fear of dogs. Such a motive could compel Agnes to seek out reasons to fear dogs by focusing on their predatory instincts and their potential for harm rather than their social tendencies. At the same time, part of her acknowledges that her fear is inappropriate. However, I agree with Greenspan that this sort of explanation is "too easy" and hardly a compelling explanation for all emotion/judgment conflicts.[223] It's unlikely that all emotion/judgment conflicts originate from secret motives in this manner, though some portion of them might. Emotion/judgment conflicts probably result most of the time from discrepancies in the outputs of our various type 1 and type 2 attitude forming processes.

In Agnes's case, fear is an automatic type 1 response. Her bad experiences with dogs solidified an association between dogs and danger or bodily harm. Her encounter

[223] Greenspan, "Emotions as Evaluations," 158.

with Millie sets in motion the association which results in her attitude of fear. Her fear is in conflict with a separate type 2 reasoned judgment that Millie is not a threat is that. We can imagine the considerations that she took into account in forming her judgment: Millie has no history of hurting people and she has very few teeth left. If she turned aggressive she wouldn't likely cause significant harm. Thus, she is not a significant threat, and fear is not an appropriate response to her. The emotion/judgment conflicts persist because the reasons on offer via Agnes's judgment are not sufficient to inhibit the associations that support her conflicting emotional attitude.

This interpretation meets the constraints described in the previous section. One constraint, provided by the plausibility argument for the principle of charity, mandated that ascriptions of irrationality should be avoided unless circumstances arise in which irrationality is likely to occur. To meet this constraint, we need only show that such circumstances obtain so that an interpretation that contains irrationality is not implausible. I have provided, in this chapter, sufficient reasons to show that the circumstances in which emotion/judgment conflicts arise are just the sort of circumstances in which it is not implausible for an attitude conflict to arise. Emotion/judgment conflicts generally arise due to discrepancies in our type 1 and type 2 processing outputs.

Another constraint—the first of three constraints supported by interpretationism—mandated that one should not attribute too pervasive errors and inconsistencies to a person. On the view I've given, conflicts among attitudes, including emotion/judgment conflicts, are somewhat common. Nevertheless, emotion/judgment conflicts are not, in typical persons, so pervasive that they would undermine the entire

foundation of interpretation for those persons. Any conflicts by their nature will be impediments to clear and determinate interpretation. However, these impediments can be great or small depending on whether they are small or large conflicts and whether they are contained within a narrow or wide domain. As Davidson suggests, it is thus a matter of degree.[224] If a person is especially prone to emotion/judgment conflicts in some domain—as Agnes is with dogs—it will be more difficult to interpret her attitudes in that domain. Yet in other domains, Agnes' speech and behavior will feature fewer conflicts and we'll have no trouble attributing a set of consistent and well-integrated attitudes to her.

The second constraint on attitude interpretation mandated that one should not attribute an attitude to someone if that attitude is not embedded in a network of attitudes and commitments that make it meaningful. On my interpretation, emotions are not unconnected to a person's attitudes. The attitudes that support them are simply separate from the ones that support the conflicting judgment. Agnes's emotion is supported by reasons settled on within the structure of implicit attitudes that yields her system 1 responses. Her judgment is supported by reasons consciously acknowledged and weighed. When conflicts occur among her attitudes, these reason sets simply come apart.

The third constraint mandated that one should not attribute inconsistencies and errors that result from systematic defects in a person's reasoning capability. My interpretation of emotion/judgment conflicts does not suggest that emotions/judgment conflicts are due to reasoning errors. Emotions that conflict with judgments need not

[224] Davidson, *Problems of Rationality*, 184.

interfere with our ability to engage in reasoning that follows rational rules using type 2 reasoning processes. Conflicts nevertheless occur because type 1 processes are not inevitably inhibited by the conflicting outputs of type 2 reasoning. It is in many ways an advantage that they are not. Since type 1 processes are capable of processing more information faster that type 2 processes, which rely on working memory, sometimes our type 1 attitudes might even be superior on a variety of metrics.[225] We must monitor for conflicts between the outputs of type 1 and type 2 processes and adjudicate between them to our best advantage.

Emotion/judgment conflicts can be difficult to resolve, but we are not without means of doing so. We can employ forms of reflection that can determine whether the emotion is appropriate and the judgment inappropriate, or whether the judgment is appropriate and the emotion inappropriate. For instance, Agnes can intentionally focus on any features of Millie's that could draw out different, less frightening associations. By this exercise, she might successfully allay her fear if there are truly compelling reasons not to be afraid to be had in the case.[226].

Since attributions of irrationality in cases of emotion/judgment conflict meet all of these constraints on such attributions stemming from the principle of charity, I argue, there is no obvious reason why we shouldn't make them. Attributions of conflict in these cases, moreover, avoid troubling implications of the notion that emotions are not

[225] See, for instance, Dijksterhuis and van Olden, "On the Benefits of Thinking Unconsciously," 627-631.
[226] See Philippot et al., "Cognitive Regulation of Emotion: Application to Clinical Disorders," 81, 83. I take up this matter in more detail in Chapter 3.

our attitudes, but nevertheless have power over our speech and behavior. We ought to view our emotions as attitudes that issue from our reasons and commitments.

5. Conclusion

The argument from charity tells us that we should not interpret a person who is experiencing an emotion/judgment conflict to have conflicting attitudes. To do so would violate the principle of charity, which enjoins us to prefer interpretations on which the person is rational over interpretations on which the person is irrational. I have argued that the principle of charity does not obviously support the conclusion that we should avoid irrational attitude ascriptions in cases of emotion/judgment conflict. I have shown that, on the contrary, irrational attitude ascriptions are within the boundaries of the constraints that the principle of charity places on attitude ascriptions. Furthermore, I have shown that if we accept the interpretationist commitments that support the strongest constraints on attitude ascriptions from the principle of charity, these commitments actually puts some pressure on us to attribute conflicting attitudes in cases of emotion/judgment conflict. Therefore, the argument from charity does not succeed. It fails to support the view that emotions must stand apart from our attitudes.

Chapter 7

Conclusion

Interest in the phenomenon of emotional recalcitrance and resistance to the influence of reflection and reasons significantly impacted much of the work on emotions that's been done in recent years. It has steered movement away from views on which emotions are or involve attitudes and it has given momentum to views on which our ownership over our emotions is generally diminished by our apparent lack of control over our emotions in emotion/judgment conflict cases. While it is true that we lack the sort of control necessary to be held accountable for our emotions in some portion of cases, I have argued that the possibility of recalcitrance shouldn't guide our accounts of emotions broadly. When recalcitrant emotions are taken as the central cases to account for rather than the aberrant ones, it leads us to view emotions as though they were commonly more stubborn and blind than they really are. They are, I have argued, the products of a generally reason responding faculty whose outputs are rooted in dispositions to react to reasons that we detect. Their outputs are also alterable with the use of various reflective tools so that they can be appropriately adjusted when our dispositions are inapt under the present circumstances. We do generally bear a proprietary and responsible relationship over our emotions since they are reason responding attitudes.

In accounting for emotional recalcitrance, we should take care to distinguish emotions that are truly recalcitrant from ones that are only apparently so. An emotion may appear to be recalcitrant because it conflicts with the deliverances of one's judgment faculty, and it seems stubborn. Emotions that appear this way may yet be

responsive to various activities of reflection on our reasons, and in some cases reflection will reveal that they may not need any revision at all. Our judgment faculty sometimes delivers attitudes that are ill-conceived, and emotion/judgment conflicts that result from the formation of such a judgment may in such cases be correctly resolved in favor of the emotion. Reflection on reasons that support having or revising each attitude involved in the conflict should move one to respond to an appropriately inclusive set of reasons with one's reason responding faculties and, consequently, it should move one to adjust one or both of the conflicting attitudes correspondingly. The attitude modifications that result might favor either the emotion or the judgment.

The often cited reasons for rejecting the possibility that emotions are anything like attitudes (due to the implications such a view would have for emotion/judgment conflicts) rests on two challenges that, I have shown, can be fully met. Since the attitude view postulates that emotion are assent-bearing attitudes, the first challenge was to show that emotion/judgment conflicts are substantially different from conflicts between two judgments. Direct conflicts between simultaneous judgments are hardly conceivable, but emotion/judgment conflicts are perfectly commonplace. It has been objected that any cognitive attitude view cannot account for this. To the contrary, I argued that the cognitive attitude view implies that there are a number of conflict possibilities that fall on a spectrum with respect to their degree of irrationality. Emotion/judgment conflicts generally fall within an acceptably ordinary range on this spectrum, with conflicts involving elements (emotions and judgments) with passive or active components that, in combination, are importantly dissimilar from the inconceivably direct conflicts of active and simultaneous judgment.

Since the attitude view implies that emotion/judgment conflicts are attitude conflicts, the second challenge to the view was to show that such attributions of irrationality wouldn't violate the principle of charity. When followed, the principle of charity preserves the plausibility of an interpretation, and it safeguards our ability to make interpretations generally by holding constant some foundational assumptions about attitudes and the agents who hold them. I argued that, while the principle of charity does place reasonable constraints on ascriptions of irrationality, interpreting emotion/judgment conflicts as irrational attitude conflicts falls within these constraints. We can preserve plausibility and our ability to interpret an agent's attitudes generally while treating emotion/judgment conflicts as the irrational, disruptive attitude conflicts that they are. To treat them as anything less would be to underplay the role that our emotions play in our lives as the deliverances of a reason-responding faculty that drives some portion of our thoughts, activities, and plans.

Bibliography

Anscombe, G. E. M. 1967. *Intention.* Ithica, N.Y.: Cornell University Press.

Arpaly, Nomi. 2000. "On Acting Rationally Against One's Best Judgment." *Ethics* 110.3: 488-513.

Audi, Robert. 1990. "Weakness of Will and Rational Action." *Austalasian Journal of Philosophy* 68.3: 279-281.

Bargh, John A., and Tanya L. Chartrand. 1999. "The Unbearable Automaticity of Being." *American Psychologist* 54.7: 462-479.

Bechara, Antoine. 2004. "A Neural View of the Regulation of Complex Cognitive Functions by Emotion." In *The Regulation of Emotion*, edited by Pierre Philippot and Robert S. Feldman, 3-32. Mahwah, NJ: Lawrence Erlbaum.

Bortolotti, Lisa. 2009. "The Epistemic Benefits of Reason Giving." *Theory and Psychology* 19.5: 624-645.

Brady, Michael S. 2014. "Emotion, Attention, and the Nature of Value." In *Emotion and Value*, edited by Sabine Roeser and Cain Todd, 52-71. Oxford: Oxford University Press.

—. 2009. "The Irrationality of Recalcitrant Emotions." *Philosophical Studies* 145.3: 413-430.

Bratman, Michael E. 1987. *Intention, Plans, and Practical Reason.* Cambridge, MA: Harvard University Press.

Child, William. 1994. *Causality, Interpretation, and the Mind.* Oxford: Oxford University Press.

Cohen, L. Johnathan. 1981. "Can Human Irrationality be Experimentally Demonstrated?" *Behavioral and Brain Sciences* 4.3: 317-370.

Cosmides, Leda. 2009. "The Logic of Social Exchange: Has Natural Selection Shaped How Humans Reason? Studies with the Wason Selection Task." *Cognition* 31: 187-276.

Cosmides, Leda, and John Tooby. 2000. "The Cognitive Neuroscience of Social Reasoning." In *The New Cognitive Neuroscience,* edited by Michael S. Gazzaniga, 1259-1270. Cambridge, MA: MIT Press.

Crandall, B., and R. Calderwood. 1997. *Clinical Assessment Skills of Experienced Neonatal Intensive Care Nurses (Final Report).* Prepared Under Contract 1 R43 NR0191101 for the National Center for Nursing, NIH, Yellow Springs, OH: Klein Associates Inc.

Damasio, Antonio. 1994. *Descartes' Error: Emotion, Reason, and the Human Brain.* New York: Putnam.

D'Arms, Justin, and Daniel Jacobson. 2003. "The Significance of Recalcitrant Emotion." In *Philosophy and the Emotions*, edited by Anthony Hatzimoysis, 129-130. Cambridge, UK: Cambridge University Press.

Davidson, Donald. 1970. "How is Weakness of the Will Possible?" In *Moral Concepts,* edited by Joel Feinberg, 93-113. Oxford: Oxford University Press.

—. 1976. "Hume's Cognitive Theory of Pride." *The Journal of Philosophy* 73.19: 744-757.

—. 1985. "Incoherence and Irrationality." *Dialectica* 39.4: 345-354.

—. 2004. *Problems of Rationality.* Oxford: Oxford University Press.

—. 1973. "Radical Interpretation." *Dialectica* 27.1: 314-328.

—. 1982. "Rational Animals." *Dialectica* 36.4: 314-328.

Davidson, Richard J., Andrew Fox, and Ned H. Kalin. 2007. "Neural Bases of Emotion Regulation in Nonhuman Primates and Humans." In *Handbook of Emotion Regulation*, edited by James J. Gross, 47-68. New York: Guilford Press.

De Houwer, Jan, and Dirk Hermans. 2010. "Do Feelings Have a Mind of Their Own?" In *Cognition and Emotion: Reviews of Current Research and Theories*, edited by Jan De Houwer and Dirk Hermans, 38-65. New York: Psychology Press.

Dennett, Daniel C. 1987. *The Intentional Stance.* Cambridge, MA: MIT Press.

Deonna, Julien, and Fabrice Teroni. 2015. "Emotions as Attitudes." *Dialectica* 69.3: 293-311.

—. 2014. "In What Sense are Emotions Evaluations?" In *Emotion and Value*, edited by Sabine Roeser and Cain Todd, 15-31. Oxford: Oxford University Press.

Devos, Thierry. 2008. "Implicit Attitudes 101: Theoretical and Empirical Insights." In *Attitudes and Attitude Change,* edited by William D. Crano and Radmila Prislin, 61-84. New York: Psychology Press.

Dijksterhuis, Ap, and Zeger van Olden. 2006. "On the Benefits of Thinking Unconsciously: Unconscious Thought Can Increase Post-Choice Satisfaction." *Journal of Experimental Social Psychology* 42: 627-631.

Döring, Sabine A. 2009. "The Logic of Emotional Experience: Noninferentiality and the Problem of Conflict without Contradiction." *Emotion Review* 1.3: 240-247.

Ekman, Paul. 1992. "An Argument for Basic Emotion." *Cognition and Emotion* 6.3/4: 169-200.

Elgin, Catherine Z. 1996. *Considered Judgment.* Princeton: Princeton University Press.

Evans, Jonathan St. B. T. 2009. "How Many Dual-Process Theories Do We Need? One, Two, or Many?" In *In Two Minds: Dual Processes and Beyond*, edited by Jonathan St. B. T. Evans and Keith Frankish, 33-54. Oxford: Oxford University Press.

—. 2007. "On the Resolution of Conflict in Dual Process Theories of Reasoning." *Thinking and Reasoning* 13.4: 321-339.

Fitzsimons, Gráinne M., and John A. Bargh. 2011. "Automatic Self-Regulation." In *Handbook of Self-Regulation: Research, Theory, and Applications*, edited by Kathleen D. Vohs and Roy F. Baumeister, 151-170. New York: Guilford Press.

—. 2003. "Thinking of You: Nonconscious Pursuit of Interpersonal Goals Associated with Relationship Partners." *Journal of Personality and Social Psychology* 84.1: 148-164.

Frankfurt, Harry G. 1988. *The Importance of What We Care About.* Cambridge: Cambridge University Press.

Frijda, Nico H. 1987. "Emotion, Cognitive Structure, and Action Tendency." *Cognition and Emotion* 1.2: 115-143.

—. 1993. "The Place of Appraisal in Emotion." *Cognition and Emotion* 7.3/4: 357-387.

Furtack, Rick Anthony. 2010. "Emotion, the Body, and the Cognitive." *Philosophical Explorations* 13.1: 51-64.

Goldie, Peter. 2009. "Getting Feelinings into Emotional Experience in the Right Way." *Emotion Review* 1.3: 232-239.

—. 2002. *The Emotions: A Philosophical Exploration.* Oxford: Oxford University Press.

—. 2006. "Wollheim on Emotion and Imagination." *Philosophical Studies* 127: 1-17.

Greenspan, Patricia S. 1988. *Emotions and Reasons: An Inquiry into Emotional Justification.* New York: Routledge.

—. 1981. "Emotions as Evaluations." *Pacific Philosophical Quarterly* 62: 158-169.

Gross, James J. 2007. "Emotion Regulation: Conceptual Foundations." In *Handbook of Emotion Regulation,* edited by James J. Gross, 3-24. New York: Guilford Press.

Gyurak, Anett, James J. Gross, and Amit Etkin. 2011. "Explicit and Implicit Emotion Regulation: a Dual Process Framework." *Cognition and Emotion* 400-412.

Haidt, Jonathan. 2001. "The Emotional Dog and its Rational Tail: A Social Intuitionist Approach to Moral Judgment." *Psychological Review* 108.4: 819, 829.

Heil, John. 1989. "Minds Divided." *Mind* 98.392: 571-583.

Helm, Bennett W. 2001. *Emotional Reason: Deliberation, Motivation, and the Nature of Value.* Cambridge: Cambridge University Press.

Henderson, David K. 1990. "An Empirical Basis for Charity in Interpretation." *Erkenntnis* 32.1: 83-103.

Hieronymi, Pamela. 2006. "Controlling Attitudes." *Pacific Philosophical Quarterly* 87: 45-74.

—. 2008. "Reasons for Believing." *Synthese* 161.3: 357-373.

—. 2005. "The Wrong Kind of Reason." *Journal of Philosophy* 102.9: 437-457.

Hill, Thomas E. Jr. 1986. "Weakness of Will and Character." *Philosophical Topics* 14.2: 93-115.

Houlton, Richard. 1999. "Intention and Weakness of Will." *The Journal of Philosophy* 96.5: 241-262.

Izard, Carroll E. 2011. "Forms and Functions of Emotions: Matters of Emotion-Cognition Interactions." *Emotion Review* 3.4: 371-378.

James, Williams. 1984. "What is an Emotion?" In *What is an Emotion? Classic Readings in Philosophical Psychology*, edited by Cheshire Calhoon and Robert C. Solomon, 127-141. Oxford: Oxford University Press.

Jones, Karen. 2003. "Emotion, Weakness of Will, and the Normative Conception of Agency." *Royal Institute of Philosophy Supplement* 52: 181-200.

—. 2006. "Quick and Smart? Modularity and the Pro-Emotion Consensus." *Canadian Journal of Philosophy* 36.32: 3-27.

Kahneman, Daniel, and Shane Frederick. 2002. "Representativeness Revisited: Attribute Substitution in Intuitive Judgment." In *Heuristics and Biases*, edited by Thomas Gilovich, Dale Griffin and Daniel Kahneman, 49-81. Cambridge: Cambridge University Press.

Kenny, Anthony. 1963. *Action, Emotion, and Will.* New York: Routledge.

Klein, Gary. 1997. "Developing Expertise in Decision Making." *Thinking and Reasoning* 3.4: 337-352.

Kolodny, Niko. 2005. "Why Be Rational." *Mind* 114.445: 509-563.

Lazarus, Richard S. 1982. "Thoughts on the Relations between Emotion and Cognition." *American Psychologist* 37.9: 1019-1924.

LeDoux, Joseph E. 2008. "Cognitive-Emotional Interactions in the Brain." *Cognition and Emotion* 3.4: 267-289.

—. 2012. "Rethinking the Emotional Brain." *Neuron* 73: 653-676.

—. 1996. *The Emotional Brain.* New York: Simon and Schuster.

—. 2013. "The Slippery Slope of Fear." *Trends in Cognitive Sciences* 17.4: 155-156.

—. 2014. "What's Basic about the Brain Mechanisms of Emotion?" *Emotion Review* 6.4: 318-320.

Leventhal, Howard, and Klaus Scherer. 1987. "The Relationship of Emotion to Cognition: A Functional Approach to Semantic Controversy." *Cognition and Emotion* 1.1: 3-28.

Loewenstein, George. 2007. "Affective Regulation and Affective Forecasting." In *Handbook of Emotion Regulation*, edited by James J. Gross, 180-203. New York: Guilford Press.

Malpus, J. E. 1988. "The Nature of Interpretive Charity." *Dialectica* 40.1: 17-35.

Mandler, George. 1975. *Mind and Emotion.* New York: Wiley.

McHugh, Connor. 2013. "Epistemic Responsibility and Doxastic Agency." *Philosophical Issues* 23: 132-157.

McIntyre, Alison. 1993. "Is Akratic Action Always Irrational?" In *Identity, Character, and Morality: Essays in Moral Psychology*, by Owen J. Flanagan and Amelie Rorty, 379-400. Cambridge, MA: MIT Press.

—. 2006. "What is Wrong with Weakness of Will?" *The Journal of Philosophy* 106.3: 284-11.

McRae, Emily. 2011. "The Cultivation of Moral Feelings and Mengzi's Method of Extension." *Philosophy East and West* 61.4: 587-608.

Millar, Alan. 2004. *Understanding People: Normativity and Rationalizing Explanation.* Oxford: Oxford University Press.

Moors, Agnes. 2014. "Flavors of Appraisal Theories of Emotion." *Emotion Review* 6.4: 303-307.

—. 2010. "Therories of Emotion Causation: A Review." In *Cognition and Emotion: Reviews of Current Research and Theories*, edited by Jan De Houwer and Dirk Hermans, 1-37. New York: Psychology Press.

Muravan, Mark, Dianne M. Tice, and Roy F. Baumeister. 1998. "Self-Control as Limited Resource: Regulatory Depletion Patterns." *Journal of Personality and Social Psychology* 74.3: 774-776.

Nisbett, Richard E., and Timothy D. Wilson. 1977. "Telling More than We Can Know: Verbal Reports on Mental Processes." *Psychological Review* 84.3: 231-259.

Nolfi, Kate. 2015. "Which Mental States are Rationally Evaluable, and Why?" *Philosophical Issues* 24: 41-63.

Nussbaum, Martha. 2001. *Upheavals of Thought: The Intelligence of Emotions.* Cambridge: Cambridge University Press.

Ochsner, Kevin N., and James J. Gross. 2007. "The Neural Architecture of Emotion Regulation." In *Handbook of Emotion Regulation*, edited by James J. Gross, 87-109. New York: Guilford Press.

Ortony, Andrew, Gerald L. Clore, and Allan Collins. 1988. *The Cognitive Structure of Emotions.* Cambridge: Cambridge University Press.

Peacocke, Christopher. 1999. *Being Known.* New York: Oxford University Press.

—. 1985. "Intention and Akrasia." In *Essays on Davidson*, edited by Bruce Vermazen and Merrill B. Hintikka, 51-73. Oxford: Oxford University Press.

Philippot, Pierre, Celine Baeyens, Celine Douilliez, and Benjamin Francart. 2004. "Cognitive Regulation of Emotions: Application to Clinical Disorders." In *The Regulation of Emotion*, edited by Pierre Philippot and Robert S. Feldman, 71-97. Mahwah, New Jersey: Lawrence Erlbaum.

Prinz, Jesse J. 2004. *Gut Reactions: A Perceptual Theory of Emotions.* Oxford: Oxford University Press.

Quine, W. V. O. 1960. *World and Object.* Cambridge, MA: MIT Press.

Raz, Joseph. 2011. *From Normativity to Responsibility.* Oxford: Oxford University Press.

Roberts, Robert C. 2003. *Emotions: An Essay in Aid of Moral Psychology.* Cambridge: Cambridge University Press.

Robinson, Jenefer. 2004. "Emotions: Biological Fact or Social Construction." In *Thinking about Feeling: Contemporary Philosophers on Emotions*, edited by Robert C. Solomon, 28-44. Oxford: Oxford University Press.

Saunders, Clare, and David E. Over. 2009. "In Two Minds about Rationality?" In *In Two Minds: Dual Processes and Beyond*, edited by Jonathan St. B. E. Evans and Keith Frankish, 317-334. Oxford: Oxford University Pres.

Saunders, Leland F. 2009. "Reason and Intuition in the Moral Life: A Dual-Process Account of Moral Justification." In *In Two Minds: Dual Processes and Beyond*, edited by Jonathan St. B. T. Evans, 335-353. Oxford: Oxford University Press.

Scanlon, T. M. 2006. *What We Owe to Each Other* . Cambridge: Belknap Press.

Schacter, Stanley, and Jerome A. Singer. 1984. "Cognitive, Social, and Psychological Determinants of an Emotional State." In *What is an Emotion? Classic Readings in Philosophical Psychology*, edited by Cheshire Calhoun and Robert C. Solomon, 173-183. Oxford: Oxford University Press.

Shaffer, Jerome A. 1983. "An Assessment of Emotion." *American Philosophical Quarterly* 20.2: 206.

Skorupski, John. 2011. *The Domain of Reasons.* Oxford: Oxford University Press.

Smith, Craig A., and Richard A. Lazarus. 1993. "Appraisal Components, Core Relational Themes, and the Emotions." *Cognition and Emotion* 7.3/4: 233-269.

Smith, Eliot R., and Elizabeth C. Collins. 2009. "Dual-Process Models: A Social Psychological Perspective." In *In Two Minds: Dual Processes and Beyond*, edited by Jonathan St. B.T. Evans and Keith Frankish, 197-212. Oxford: Oxford University Press.

Smith, Michael. 1995. *The Moral Problem.* Oxford: Blackwell.

Solomon, Robert C. 2004. "Emotions, Thoughts, and Feelings: Emotions as Engagements with the World." In *Thinking about Feeling: Contemporary Philosophers on Emotions*, by Robert C. Solomon, 76-90. Oxford: Oxford University Press.

—. 2003. *Not Passion's Slave: Emotions and Choice.* Oxford: Oxford University Press.

Stanovich, Keith E. 2009. "Distinguishing the Reflective, Algorithmic, and Autonomous Minds: Is it Time for a Tri-Process Theory?" In *In Two Minds: Dual Processes and Beyond*, edited by Jonathan St. B. T. Evans Evans and Keith Frankish, 55-88. Oxford: Oxford University Press.

—. 2004. *The Robot's Rebellion.* Chicago: University of Chicago Press.

Stein, Edward. 1996. *Without Good Reason: The Rationality Debate in Philosophy and Cognitive Science.* Oxford: Oxford University Press.

Stich, Stephen. 1985. "Could Man be an Irrational Animal." *Synthese* 64.1: 115-135.

Strawson, Galen. 2003. "Mental Ballistics or the Involuntariness of Spontaneity." *Proceedings of the Aristotelian Society* 103: 227-256.

Tappolet, Christine. forthcoming. "Emotions, Perceptions, and Emotional Illusions." In *The Crooked Oar, the Moon's Size and the Kanizsa Triangle: Essays on Perceptual Illusions*, by Clotide Calibi. Cambridge, MA: MIT Press.

Taylor, Gabriele. 1985. *Pride, Shame. and Guilt: Emotions of Self-Assessment.* Oxford Oxfordshire: Clarendon Press.

Thagard, Paul, and Richard E. Nisbett. 1983. "Rationality and Charity." *Philosophy of Science* 50.2: 250-267.

Tiberius, Valerie. 2013. "In Defense of Reflection." *Philosophical Issues* 23: 223-243.

—. "Wisdom and Perspective." *Journal of Philosophy* 102.4: 163-182.

Toribio, Josepha. 2011. "What We Do When We Judge." *Dialectica* 65.3: 345-367.

Tracy, Jessica L., and Daniel Randles. 2011. "Four Models of Basic Emotions: A Review of Ekman and Cordano, Izard, Levenson, and Panksepp and Watt." *Emotion Review* 3.4: 397-405.

Walker, Mark Thomas. 1996. "The Voluntariness of Judgment." *Inquiry* 39.1: 97-119.

Wason, P.C. 1968. "Reasoning about a Rule." *Quarterly Journal of Experimental Psychology* 20.3: 273-274.

Williams, Bernard. 1973. *Problems of the Self : Philosophical Papers 1956-1972.* Cambridge: Cambridge University Press.

Wilson, Timothy D., Jonathan D. Schooler, Sara D. Hodges, Kristen J. Klaaren, and Suzanne J. LaFleur. 1999. "Introspecting about Reasons can Reduce Post-Choice Satisfaction." *Personality and Social Psychology Bulletin* 19.3: 331-339.

Wilson, Timothy D., Samuel Lindsey, and Tonya Y. Schooler. 2000. "A Dual Model of Attitudes." *Psychological Review* 107.1: 102.

Wollheim, Richard. 1999. *On the Emotions.* New Have: Yale University Press.

Zajonc, Robert B. 1980. "Feeling and Thinking: Preferences Need No Inferences." *American Psychologist* 35.2: 151-175.

—. 1984. "On the Primacy of Affect." *American Psychologist* 39.2: 117-123.