Emotions as Reasons: Against the Standard Belief/Desire Account of Action

Dissertation

Presented in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in the Graduate School of The Ohio State University

By

Lindsey Teague Mason, B.A.
Graduate Program in Philosophy

The Ohio State University
2013

Dissertation committee:
Abraham Roth advisor
Timothy Schroeder
Justin D'Arms
Abstract

In this dissertation, I will defend the thesis that emotions are reasons for actions. I defend this thesis against the standard account of action, for which only belief/desire pairs are reasons for action. As preliminaries, I defend an embodied appraisal view of what emotions are, and I present a Burge-style entitlement view of how emotions, which are nonpropositional, can be reasons. I begin my main argument by considering purely expressive actions, those actions which, according to Rosalind Hursthouse, are intentional, not done in order to achieve some further goal, and cannot be explained without reference to emotion. In chapter 4, I argue that all attempts to save the standard account from this challenge fail. In chapter 5, I argue against two attempts to explain expressive actions outside the standard view, and then I present my own view. Then in the last chapter, I present a more broad view of how emotions rationalize actions in non-expressive ways.
I dedicate this dissertation to my beautiful daughter, Jolie, and the love of my life, Rob, for all the sacrifices they've made in this process.
I am indebted in this work, especially, to my advisor Abraham Roth. He was harder on me than anyone has ever been. He pushed me to clarify every single idea in this dissertation. He challenged me to anticipate objections from the philosophical community, especially philosophers working in action theory. And he forced me to defend my conclusion against each of those objections. I wouldn’t be here if it weren’t for him. I’d also like to thank Timothy Schroeder for advising me through the Candidacy and Qualifier exams. Your guidance, loyalty, and efforts were greatly appreciated. Justin D’Arms has also been extraordinarily helpful in clarifying many of the ideas here, in private conversations as well as subject-related seminars. Thanks to all of the other faculty, graduate students, and staff in the philosophy department for all of their hard work, invaluable feedback, and a great environment for honing my philosophical skills. This has been the time of my life.

I also note permission to other copyrighted materials used extensively in this text.
Vita

May 1999 ................................................................. Dacula High School

2005 ................................................................. B.A. Philosophy, Auburn University

2005 to present ........................................................ Graduate Teaching Associate,

                      Department of Philosophy, The Ohio State University

Fields of Study

Major field: Philosophy
Table of Contents

Abstract ............................................................................................................. ii
Dedication.......................................................................................................... iii
Acknowledgments............................................................................................... iv
Vita....................................................................................................................... v
List of Tables....................................................................................................... vii
Chapter 1: Introduction...................................................................................... 1
Chapter 2: What Emotions Are.......................................................................... 11
Chapter 3: An Entitlement Theory................................................................. 32
Chapter 4: On Actions Expressing Emotions.................................................. 53
Chapter 5: Possible Alternatives to the Standard Account............................. 82
Chapter 6: Emotions as Reasons....................................................................... 115
References......................................................................................................... 154
List of Tables

Table 1: Core Relational Themes ................................................................. 107
Table 2: Action Tendencies ................................................................. 110
Chapter 1: Introduction

Joan and Jane are best friends. One day, though, Jane finds out that her husband has been having an affair with Joan. Jane is enraged by this news! She feels so betrayed by both of them—but mostly by Joan, her life-long friend. “How could she do this to me?” Jane screams. And just then—in a wave of rage—Jane rushes over to the mantel, where a picture of smiling Joan, Jane, and her husband sits. Jane pulls the photograph down, rips the photo out of the frame, and she stares at it. “How can this be happening?” she thinks. “All of this happiness is gone, and it’s all your fault” she says, staring at Joan in the photo. She hurriedly brings the photo with her into the office, where she snatches a pair of scissors. With the scissors in hand, and from all of that anger, she scratches Joan’s eyes out of the photograph.

Merle just lost his wife. He is overcome with grief. They had been married 46 years. “What will life be like without her?” he mutters through tears and sorrow. He’s stumbling through the house, hunched, grabbing on to the walls for support. He wanders into the bedroom, and there he sees the closet. He goes and pulls out some of his late wife’s clothes. He holds them close to him, kisses them, caresses them, and sits down on the floor with them. The grief overcomes him, and he lies down, rolling around in his dead wife’s clothes.
Kate is alone in the house. Her parents have gone to dinner; her brother is spending the night at a friend’s. As Kate watches a movie, she hears a loud bang—it seems to be coming from the kitchen. Kate’s heart begins to race, her palms sweat, she can hear someone trying to break into her house! She grabs her cell phone and runs upstairs. She calls 911. Then—out of intense fear—she hides her face under the blankets of her bed.

Jane, Merle, and Kate are all acting expressively. Jane expresses her anger by scratching Joan’s eyes out of the photograph; Merle expresses his grief by rolling around in his late wife’s clothes; Kate expresses her fear by hiding under her bedspread. Their actions all express emotions. Philosopher Rosalind Hursthouse considers actions such as these in her paper “Arational Actions.”¹ What is so interesting about these actions? Well, as Hursthouse first noticed, expressive actions seem to challenge a long-held belief about intentional action. Most contemporary philosophers working in action theory (and many working in other areas as well) have taken a certain view of intentional action as the only viable option. There has been such agreement on this “standard account” that few challenge it’s truth. I believe—with Hursthouse—that expressive actions do challenge this standard account of action.

But I should be clear about what I mean by “expressive actions” and what I mean by the “standard account.” I’ll take Hursthouse’s canonical paper as the

¹ What Hursthouse calls “arational actions” I will call “expressive actions” throughout this dissertation. I think “expressive actions” is a better term because it leaves it up for debate whether such actions are rational or not.
standard for defining expressive actions. According to Hursthouse, of actions such as Jane's, Merle's, and Kate's (among others), it would be true to say:

(i) that the action was intentional;
(ii) that the agent did not do it for a reason in the sense that there is a true description of action of the form “X did it (in order to) ...” or “X was trying to...” which will “reveal the favorable light in which the agent saw what he did”, and hence involve, or imply, the ascription of a suitable belief; and
(iii) that the agent would not have done the action if she had not been in the grip of whatever emotion it was, and the mere fact that she was in its grip explains the action as much as anything else does (Hursthouse, 59).

Condition (i) is supposed to contrast cases of accidents and mere bodily movement—movements such as clenching one’s fists in anger or jerking around in surprise—with the case of purposeful, directed actions. When Jane scratches the eyes out of the photograph, this is not a mere bodily movement. She must direct the scissors in very clear ways to achieve the goal of scratching only Joan’s eyes. If she also scratched the table underneath the photograph, this would be an accident, and non-intentional. But scratching the eyes in the photograph is intentional.²

Condition (ii) is making reference to beliefs which are about the means for action. According to the standard account (simply put for now), in order for an action to be properly rationalized, the agent must have identified some goal for action, and the agent must have some belief about how to achieve that goal. Such beliefs are “means-end beliefs” because they are about the means for achieving the end set by some other state, usually desire. But such means-end beliefs seem to be

² Some have disputed this claim, and I will address those arguments in chapter 4.
absent for Jane, Merle, and Kate. Jane is not scratching the eyes out of the photograph as a means to achieving some further end—say, harming the real Joan. She’s just scratching the eyes up. Merle isn’t rolling around in his dead wife’s clothes as a means for bringing her back or for even remembering her. He’s just rolling around in them. And Kate is not hiding under her bedspread as a means to escaping the intruder. She probably doesn’t even believe that she’s hiding well—she’s just hiding out of fear.\(^3\)

Condition (iii) references the emotions Jane, Merle, Kate and the like are expressing. Jane wouldn’t be scratching out the eyes if she weren’t angry. Merle wouldn’t be rolling around in his deceased wife’s clothes if he weren’t grieving. And Kate wouldn’t hide under her bedspread if she weren’t afraid. Moreover, if we were to walk in on Jane, for instance, and ask, “what are you doing in here?” the answer we’ll likely get is “I’m so angry with Joan!” The agents take their own emotions to explain their actions, and when we discover what their emotions are, we think they explain the agent’s actions as well.

So where does this leave the standard account of action? Here’s my definition of the standard account:

**Standard Account:** For all actions, A’s action is intentional if and only if she does it for a reason where performing the action for a reason requires (1) A has some element in her subjective motivational set (usually desire) the satisfaction of which will be served by her φ-ing, (2) A has a belief about the

---

\(^3\) There are several other ways of formulating the requisite belief, which I will discuss in much more detail in chapter 4.
means for satisfying her desire, and (3) A’s action involves a practical inference and not mere content-efficacy.\textsuperscript{4}

Ignoring condition (3) for now, the standard account requires “belief/desire pairs.” It requires the agent desire X and believe that performing action A will bring about X. In my examples above, the agents simply do not have such belief/desire pairs. Jane doesn’t desire to harm Joan and believe that scratching the eyes out of the photograph will harm her. That would be a ridiculous belief! But even if the agent does have a requisite belief/desire pair (suppose Jane desires to scratch the eyes out of the photograph, and she believes that doing this-and-that with the scissors will scratch the eyes out) the belief/desire pair does not \textit{satisfactorily} explain Jane’s action. If we walked into the room and saw Jane there, asked her why she’s doing that, and she replied “because I wanted to scratch these eyes out, and I believed that doing this-and-that with the scissors would do that” we would be very unsatisfied. Instead, we’d have to hear some reference to her anger with Joan to be satisfied. And so, it seems the standard account as it’s often articulated is inadequate because it cannot make sense of expressive actions.\textsuperscript{5}

Instead of the standard account, I propose a different account of action that addresses those actions performed by Jane, Merle, and Kate. My account relies on the emotions:

\textbf{Emotion-Based Rationalizations:} For some actions, A’s action is done for a reason if and only if (1) A has some element in her subjective motivational

\textsuperscript{4} More on the difference between practical inference and content-efficacy will follow in chapter 6.

\textsuperscript{5} These arguments will be repeated and further developed in chapters 4, 5, and 6.
set (usually desire) the satisfaction of which will be served by her ϕ-ing. (2) A has an emotion that represents a relation between herself and something she values in the world, and (3) A’s action falls within the normal range of action tendencies for the relevant emotion.

The big implication of my view is that an agent can, at times, act for a reason by acting on an emotion/desire pair. Jane is full of anger, and this anger combines with a desire to scratch out Joan’s eyes in the photograph to rationalize her action. Merle is full of grief, and this grief combines with a desire to roll around in his dead wife’s clothes to rationalize his action. Kate is afraid of the intruder, and this fear combines with a desire to hide under the bedspread to rationalize her action. I believe this is the simplest account of actions such as Jane’s, Merle’s, and Kate’s. When we hear such an explanation, we are satisfied.6

The purpose of this dissertation is two-fold: it is an argument against the standard account, and it is an argument for the emotion-based rationalizations. The argument against the standard account relies on considering expressive actions. But a full defense of the emotion-based rationalizations will rely on expressive as well as non-expressive actions done out of emotions. What is a non-expressive action done out of emotion? Compare Kate to Kim. Kim is in the same situation as Kate, except instead of hiding under a bedspread after calling 911, Kim hides in a secret crawl space in the basement. A normal belief/desire pair could explain Kim’s action—she wants to hide from the intruder, and she believes that hiding in the

---

6 I’ll discuss condition (3) in more detail in chapter 5.
secret crawl space will best achieve that goal. In my final chapter, I argue that even
actions such as Kim’s must reference emotions to be rationalized.

Here’s a road map to the dissertation. In chapter 2, the goal is to define what
I take emotions to be. If I’m arguing that emotions can be reasons for action, then I
believe it’s necessary for me to define what emotions are. There are many different
theories of emotions on the market. I argue that any strict cognitivism that requires
a belief or judgment as essential to emotions is false. And I argue that a strict feeling
theory that equates emotions only to the physical feelings in the body is false.
Instead, emotions must be felt in the body (contrary to strict cognitivism), but they
also must be intentional states—states that are about things—contrary to strict
feeling theories. Emotions are feelings of the body, but they also represent things in
the world. This leads me to adopt Jesse Prinz’s embodied appraisal theory of
emotions from *Gut Reactions*. I then rely on that theory throughout the dissertation.

The goal of chapter 3 is to say how emotions as embodied appraisals can be
reasons. This is actually a big problem for me. According to the embodied
appraisal view, emotions are nonpropositional states. They aren’t constituted by
beliefs, judgments or anything like beliefs and judgments, and so they aren’t
constituted by propositional states. I think this must be the case so that small
children and nonhuman animals that do not possess propositional attitudes can still
experience emotions. But on most views, for a state to be a reason, it must be
propositional. I solve this problem by drawing on work by Tyler Burge, who points

---

7 And thinking of embodied appraisals as reasons diverges from Prinz’s own view.
out a parallel problem in the case of perception. Burge argues that perceptual experiences are nonpropositional, and yet we can move from such nonpropositional contents to propositional attitudes—perceptual beliefs—in a rational way. He claims that in the case of perception, we are entitled to rely on our perceptual experiences, where “entitlement” is a technical term for a specific kind of warrant. I simply take Burge’s solution for perception and apply it to the case of emotions. So, while emotions are nonpropositional, they can still be reasons for propositional beliefs and judgments, as well as actions.

Chapter 4 then turns to expressive actions and the standard account. I present in more detail how expressive actions challenge the standard account. I then spend the rest of the chapter presenting several views meant to save the standard account from this challenge. By tweaking the description of expressive actions or amending the standard account in various ways, Joseph Raz, Michael Smith, Peter Goldie, and David Velleman all try to save the standard account from the expressive challenge. I go through each of these attempts and find them inadequate. I conclude that the challenge remains.

Given that the challenge remains, in chapter 5, I consider two other accounts of expressive actions that do not assume the standard account. Monika Betzler and Sabine Doring each give an account of expressive actions as rationalized by

---

8 He actually argues for something stronger—that perceptual experiences are nonconceptual. And then he argues that we can move from such nonconceptual states to propositional attitudes in a rational way. But supposing propositions are composed of concepts, if the argument works for the big move from nonconceptual content to propositional content, then it will be easier to make it work for (potentially conceptual) nonpropositional content to propositional content.
something other than belief/desire pairs. Betzler argues they are rationalized by
the agent’s ongoing evaluative perspective, and Doring argues they are rationalized
by the emotion alone. I find serious problems with each of these views. I then
spend the second half of the chapter presenting my positive view of expressive
actions, which leads me to the emotions-based rationalization principle. I think
expressive actions challenge the standard account, and so my positive proposal is
contrary to it. I conclude we ought to abandon the standard account as an account
of all intentional action.

In the final chapter, I expand my positive proposal—the emotions-based
rationalization principle—to other actions done for emotions. I explain in some
detail what it means to act for a reason on a number of different views. I then
consider a series of examples where an agent is acting for a reason, but the earlier
views cannot capture the sense in which the agent is reason responsive. The only
view, I conclude, that can make sense of the agent’s reason responsiveness is my
emotions-based rationalization view. Then I must distinguish my view from a
variety of other views, and argue that my view is the best. Agents can act for a
reason by acting on an emotion/desire pair. At this point, I’ve taken myself to have
fully defended this conclusion.

The standard account of action is so entrenched in this discipline; it is
definitely difficult to challenge it. I have attempted to do so here. I believe that the
best challenge to the standard account is through the emotions. Human beings are
emotional creatures. We have evolved to be emotional—it’s often how we
communicate and get along with each other. We are social animals. We laugh to form bonds; we get angry to enforce rules; we cry to evoke sympathy and have our needs met; we feel guilt to induce mercy. We make decisions based on how we feel now, or on how we imagine we’d feel in the future, or even on how we felt in the past. Emotions are central to our lives. They determine much of how we get around in the world. For the standard account to ignore all of that is egregious. It’s time we start incorporating emotions into our view of human action and rationality. It’s time we start seeing that the reason human beings do very much of what they do is based on how they feel—not what they think. And there’s nothing wrong with that.
Chapter 2: What Emotions Are

I. Introduction

York Gunther argues “Any satisfactory model of the emotions must at once recognize their place within intentional psychology and acknowledge their uniqueness as mental states” (Gunther 279). I take this to be an adequacy constraint on any theory of what emotions are. Emotions must be (1) mental states with intentional content, i.e., they must be about something, and (2) unique mental states i.e., they must have a distinctive functional role.

Some emotion theorists have denied one or the other of these constraints. For instance, if we give up the requirement that emotions are mental states with intentional content, then we get the feeling theory of William James. The feeling theory says that emotions are feelings of changes in the autonomic and motor functions of one’s body. When we perceive that we are in danger, for example, this perception sets off a collection of physiological responses, and these bodily feelings are what constitute fear. These feelings are similar to pain, for instance, in that

---

they’re not “about” anything. This theory is often attributed to William James historically, and Antonio Damasio contemporarily.

Suppose instead a theorist denies the requirement that emotions are unique mental states. An example of this move can be found in the strict cognitivist tradition.\(^\text{10}\) Such theorists give up the uniqueness by reducing emotions to judgments or beliefs. For instance, a pure cognitive theory says emotions are identical to thoughts. Fear just is the thought that such-and-such is dangerous. One isn’t feeling guilt without the thought that one has done something wrong. Historically, this theory dates back to the Stoics. But more recent defenders include Robert Solomon and Martha Nussbaum.

A composite cognitive theory—one that involves more than one kind of propositional attitude—would also violate the uniqueness condition. Such a theory can be found in O.H. Green. He argues that there are four basic emotions, gladness, hope, sorrow, and fear (Green, 83). These cardinal emotions are all reduced to a belief/desire pair. Gladness is a belief with certainty that p and a desire that p. Hope is a belief without certainty that p and a desire that p. Sorrow is a belief with certainty that p and a desire that not-p. And fear is a belief without certainty that p.

---

\(^{10}\) There is a lot of debate about how to define “cognitivism.” Here I use the terms “strict” cognitivism or “pure” cognitivism to refer to a subset of cognitive views where emotions are identical to judgments or beliefs. I will be arguing against two such cognitivist theories: judgmentalism and Green’s belief/desire account. I do not take these to be the only cognitivist theories. Some appraisal theories are cognitivist in the sense that emotions require judgments, but the emotion is not identical to judgments. Other cognitivist theories require emotions to be conceptual, but not propositional—if you want to call such a theory “cognitive” at all. I will just stipulate by “strict cognitive” I mean the theory that emotions are identical to evaluative judgments or the theory that emotions are identical to belief/desire pairs.
and a desire that not-p. Emotions are not unique if they are simply reducible to judgments, beliefs, thoughts, desires or some combination of those states. Strict cognitivism, then, is either the view that emotions are identical to judgments or beliefs, or it is the view that emotions are identical to belief/desire pairs.

The position I will defend in this paper is that emotions are intentional representational states which are not reduced to any other intentional state. I want to guarantee both intentionality and uniqueness. I want to satisfy the adequacy constraint proposed by Gunther.

How will I do that? First, I’ll argue against cognitivism (both Solomon and Green). Here, the main line of attack is that we can have emotional experiences in the absence of judgments. There must be uniqueness to emotions. Second, I will argue against the feeling theory. While feeling is essential to an emotional experience, there seems to be more to pride than the swelling of one’s chest. There seems to be intentional content. Third, I’ll introduce a specific kind of theory—the embodied appraisal theory—which I will show meets the two adequacy constraints. I will explain and adopt Jesse Prinz’s version of appraisal theory with a few adjustments.11 The goal of this chapter is quite narrow. I hope to give some intuitive reasons for rejecting strict cognitivism about the emotions as well as a strict feeling theory. Instead, I think we ought to accept a theory of the emotions which encompasses both feelings and appraisals of the world. Prinz’s theory does

11 The purpose of this chapter is not to thoroughly defend Prinz’s theory as the best overall account of emotions. While I take it to be the best account, and I will use this account throughout this dissertation, the account is not without flaws.
just that. And I will be relying on his appraisal theory throughout this dissertation. So, here I’m simply laying out what I take emotions to be such that they can be reasons for action—the goal of the dissertation.

II. Argument Against Strict Cognitivism

Robert Solomon argues that emotions are evaluative judgments (henceforth called “judgmentalism”). He claims, “I want to follow the Stoics and Spinoza in arguing that emotions are judgments, much more akin to thoughts than to physiological or physical commotion” (Solomon, 94). While it’s unclear to me exactly what Solomon means by “judgment,” I will suggest an interpretation that makes the most sense. The first thing to note is that all emotions are constituted by evaluative judgments. That is, without such a judgment of the form ‘flying is dangerous’, one would not be experiencing fear of flying. Second, emotions are constituted only by evaluative judgments. The physiological changes we feel when we are afraid—a racing heartbeat, sweating, and a feeling of dizziness—are not essential to fear itself. We could, in principle, be feeling fear in the absence of such bodily changes.13

---

12 I use ‘cognitivism’ here to be a broad term covering Solomon’s judgmentalism and Green’s reduction of emotions to belief/desire pairs. There are other emotion theories that might qualify as cognitive but not as judgmentalist, for instance—such as some versions of appraisal theories. I will not be addressing these other cognitive theories here.

13 Solomon argues that such judgments are intimately connected to desires, and this helps to make sense of the “arousal” we feel. However, I’m not questioning the cognitivist’s rejection of feeling; I am questioning the idea that judgments are essential to emotional experience.
I believe that emotions are not evaluative judgments. That is, evaluative judgments are not constitutive of emotions, and bodily changes are essential to being in an emotional state. I’ll take each of these in turn.

The first problem for judgmentalism is that judgments are not necessary for feeling an emotion. The judgmentalist is committed to judgments being necessary for emotional experience. A consequence of such a commitment is that someone who does not believe that flying, for instance, is particularly dangerous cannot be genuinely afraid of flying. This is problematic because it seems people do feel fear without such a judgment. One might experience fear of flying in the absence of a judgment “flying is dangerous.” The aviophobic might feel fear—she might feel the tightness in her chest, her hands gripping the arm rests, the sweat dripping down her temple, the hairs lifted on the back of her neck—without ever forming the judgment that flying is dangerous. And more importantly, one might experience fear of flying while judging that flying is perfectly safe. In fact, this is what happens in most cases when people have a phobia—such as an irrational fear of flying. The person boarding the plane, waiting for takeoff, repeatedly says to herself, “flying is not dangerous; flying is perfectly safe; driving to Florida would be statistically more dangerous than flying.” And yet the fear persists. The fear persists in the absence of a judgment of the form “flying is dangerous” and in the presence of a judgment to
the contrary. If so, then evaluative judgments are not required for being in an emotional state. And so, they are not constitutive of emotions.

Sometimes judgments are part of an emotional experience. I might not feel jealous, on occasion, without the thought that my partner has betrayed me. But there are also cases where I feel jealous without such a judgment. Suppose I discover my partner having lunch with a beautiful stranger—something I didn’t know was going to happen. As I approach to say hello, I’m imagining he’s having an affair with this woman, and I’m growing more jealous with each step. When I reach the table, he sees me, jumps up and introduces his sister (such a scene has been reproduced in movie after movie). While I do feel an immediate sense of relief, and I now know he’s not having an affair at all, that feeling of jealousy lingers, still, for a while. It doesn’t automatically dissipate once I judge that my partner’s affections are still all mine. This same lingering occurs with many strong emotions—your fear might not automatically subside when you escape the hairy spider. Your anger continues sometimes days after you discover your “assailant” really meant no harm. This lingering effect of emotions seems to suggest they can persist far beyond the reaches of an evaluative judgment.

The second problem for judgmentalism is that it cannot satisfy one of the two constraints on an adequate emotion theory. Emotions must be unique mental

---

14 See D’Arms and Jacobson, “The Significance of Recalcitrant Emotion (or, Anti-Quasijudgmentalism)” for a more detailed argument against judgmentalism along these lines. I don’t have the space to completely lay out such an argument here—I hope the possibility of emotional experience absence a judgment is intuitive enough for my purposes.
states. I’ll take our language here to be a clue. If we could do everything we wanted to do with only judgments, then why do we need the word ‘emotion’? Presumably emotions are distinct from judgments—they are not simply a special species of judgment. They need to have a unique functional role. Emotions seem to have developed to help us survive and adapt to changing circumstances in a way judgments did not. The fight or flight response to a dangerous stimulus is certainly necessary for survival. Guilt helps us survive by diminishing others’ anger and keeping the peace in a society. And so on for many of our basic emotional responses. Emotions serve a very important function in the lives of human beings, and I think we’d be running roughshod over that unique functional role by simply assimilating them to some other mental state. And so, I think it’s too simplistic to posit that emotions are simply judgments. Judgments cannot do all the functional work emotions can do.\textsuperscript{15}

The third problem with cognitivism concerns the absence of feeling. What would it be like to be in the grip of sadness without the tears and heavy weight on one’s chest? That is, is one really sad without feeling the physiological changes that seem to accompany sadness? I must admit that it’s hard for me to imagine. But maybe that’s not exactly what the proponent of emotions being constituted by

\textsuperscript{15} This is a rough, intuitive argument that judgmentalism is inadequate. Someone such as Solomon will likely respond by incorporating all of the elements I mention above into the concept of a judgment. But then it is hard to track what we mean by “judgment” anymore. There is a long history of debate over cognitivism, and I could write an entire dissertation on that topic alone. Here all I want to establish is that there is some reason to believe an adequate constraint on a theory of emotions is that they be unique mental states. And reduction of emotions to other states such as judgments will violate this constraint.
evaluative judgments is willing to accept. Rather, the evaluative judgments cause the physiological changes. But the physiological changes are not constitutive of being in an emotional state.

I don't think causation is enough. So long as bodily changes are not constitutive of emotional states, it is in principle possible to be sad without feeling tears well up or feeling the heartbeat slow or feeling a weight on one’s chest. I don’t think this is a conceptual possibility. We do feel emotions. We even use the terms “feelings” and “emotions” as interchangeable quite often. Feeling emotions is essential to being in an emotional state. Otherwise, there’s nothing distinctive about emotions. They would simply be like any other evaluative judgment such as “lying is wrong.” But there is something intuitively distinctive about emotions—and this is that they are felt in the body. And so, feeling an emotion must be constitutive of being in an emotional state.

While I’ve argued that evaluative judgments are not constitutive of emotions, that doesn't mean such judgments have no role to play. Sometimes we feel emotions in response to our evaluative judgments. And at other times, we might form an evaluative judgment in response to an emotional state. If I believe or judge that someone has offended me, I am more likely to then feel anger towards the person. Likewise, if I feel anger well up inside me, I am more likely to judge that someone has offended me. We often feel emotions such as guilt at the same time we judge that we’ve done something wrong. They occur simultaneously. And this can persuade us to think they are the same states, surely. But when it comes to theories,
we ought to keep a metaphysical account of what emotions are separate from an account of what judgments are.

Another cognitive view of emotions might be that they are belief/desire pairs. O.H. Green accepts such a view. He argues that there are four basic emotions, gladness, hope, sorrow, and fear (Green, 83). Gladness is a belief with certainty that p and a desire that p. Hope is a belief without certainty that p and a desire that p. Sorrow is a belief with certainty that p and a desire that not-p. And fear is a belief without certainty that p and a desire that not-p.

However, Green's view appears to come at a cost. Many philosophers and psychologists (James, Lange, Damasio, Griffiths, Prinz, and the list goes on) have argued that emotions involve essentially some physiological changes in the body. This is often termed “affect,” and includes changes in the respiratory system, the musculoskeletal system, the circulatory system, the endocrine system, etc. It is up for debate whether these changes need to be felt or merely present. But it seems highly intuitive that what makes a given experience distinctly emotional is the changes we experience in our bodies. The racing of the heart in fear, the weeping sluggish feeling in sorrow, and the weight on the shoulders in guilt are important features of our emotional lives. Green seems to have ignored these bodily changes just as other cognitive theorists have.

Well, Green hasn't completely ignored affectivity. He has a response available. He argues that not only are emotions belief-desire structures, they are *hedonic* belief-desire structures. The congruent emotions (ones that involve both a
belief that p and a desire that p) are pleasurable. The divergent emotions (ones that involve a belief that p, but a desire that not-p) are unpleasant. It is the hedonic character of emotions that give us affect. We feel pleasure and displeasure in our bodies.

I do not think that pleasure and pain alone can account for all of the bodily changes that occur in an emotional experience. Pleasure and pain merely tell us whether we are faring well or faring poorly in the world. Emotions, however, can tell us many different ways of faring poorly, at least. The displeasure felt in sadness is very different from the displeasure felt in fear. Sure, they are both forms of displeasure. But Green cannot account for why or how sadness feels different from fear.

But there is a further problem with the class of experiential emotions. Imagine that I perceive my cup of coffee on my desk and desire my coffee. This is a perception-desire structure, but is it an emotion? No, surely it is not. But by arguing both that all hedonic states are emotions and that perceptions can combine with desires over a shared topic to form emotions, Green opens up the class of emotions much too broadly. And so, we ought to reject Green’s belief/desire model for emotions as well as Solomon’s claim that emotions are judgments. Both fail the adequacy constraint that emotions be unique mental states. Emotions for the strict cognitivist are reduced to judgments or belief/desire pairs. So, we need to look elsewhere for an adequate theory.
III. Argument Against the Feeling Theory

If the problem with cognitivism is not enough feeling, then we might retreat to a pure feeling theory, such as William James’. James argues that many theorists believe that there’s some “mental perception of some fact” which causes the emotion which then gives rise to a bodily expression (i.e., smiling). James, however, wants to reverse things. He argues, “My theory, on the contrary, is that the bodily changes follow directly the perception of the exciting fact, and that our feeling of the same changes as they occur IS the emotion” (James, 449). And so, “we feel sorry because we cry, angry because we strike, afraid because we tremble, and not that we cry, strike, or tremble, because we are sorry, angry, or fearful” (James, 450). So, the feeling theory says: an emotion is identical to the feeling of bodily changes that arise from some external elicitor.

James has a pretty good reason for thinking such feelings of bodily changes are essential to emotions. He argues,

If we fancy some strong emotion, and then try to abstract from our consciousness of it all the feelings of its bodily symptoms, we find we have nothing left behind, no ‘mind-stuff’ out of which the emotion can be constituted, and that a cold and neutral state of intellectual perception is all that remains (James, 451).

Can you imagine what it would be like to be afraid—genuinely afraid—of a charging bull without feeling, as Darwin notes, “the heart beat[ing] quickly and violently, so that it palpitates or knocks against the ribs” (Darwin, 290-292)? Would you even be in a state of fear without a trembling of all the muscles in one’s body or the pallor of
one’s face? Can one feel fear without the hairs rising all over one’s body? What would fear be like without any such physical manifestations? James claims, “A purely disembodied human emotion is a nonentity” (James, 452). I tend to agree.

However, we must be careful of a relevant distinction here. One might say that emotions are \textit{identical} only to feelings of bodily changes (as James does), or one might say that feelings of bodily changes are \textit{essential} to emotions without being strictly identical. The latter option leaves open the possibility that there is more to an emotion than feelings of bodily changes. James’ argument from introspection above would be consistent with either option. Which should one accept?

If we are to follow the adequacy constraints, then we ought to accept the latter because only it (might) involve intentional mental states. If emotions are identical to feelings of bodily changes, and those feelings are (as James supposes) nonintentional, then we've violated the constraint. This is some reason to accept the second option.

But there is a stronger reason to accept the second option, which arises from the way we individuate emotions. When saying I felt fear rather than anger, I do not cite the trembling of my hands versus the clenching of my fists. I do not use the feelings of bodily changes to differentiate between various emotions. I don’t even think I could. Instead, I say I felt afraid of the \textit{dangerous} thing and angry at the \textit{offense}. We seem to differentiate emotions by their elicitor rather than by feelings of bodily changes. So, while such feelings may be essential to emotions, so too may they be a special kind of object that reliably elicits specific emotions.
According to Kenny (1963) emotions can have two kinds of objects. A *formal object* is the property in virtue of which an event elicits an emotion. A *particular object* is the event itself. Take fear for example. The formal object of fear is the property of being dangerous, but a charging bull is the particular object of an instance of fear. Fear—which essentially involves feelings of bodily changes—is set off by dangerous things in general, and is about the charging bull right now. The only way to achieve intentionality is to give fear an object—Kenny gives fear two kinds of objects, one parasitic on the other. And it’s only through this intentionality that we can distinguish fear, with the property of being dangerous as its formal object, from anger, with the property of being an offense as its formal object.

So, while feelings of bodily changes are essential to emotions, we must also have a better way of differentiating one emotion from another. This goal is achieved by giving emotions formal objects. And then we achieve the intentionality of emotions by giving emotions specific content: instantiations of the formal objects with particular objects. How does this work? In the next section I will present Prinz’s theory of the emotions. He argues that emotions are “embodied appraisals.” Emotions are embodied to account for the feelings of bodily changes, and emotions are also appraisals of things in the world to account for the intentionality of emotions. I argue that Prinz’s theory properly satisfies the adequacy constraints above.
IV. Prinz’s Embodied Appraisals

Prinz’s theory is part Jamesian and part appraisal theory. He agrees that bodily changes are essential to emotions, but he thinks emotions also represent something external to one’s body. That is, emotions appraise what matters to us in the world. Emotions represent relational properties such as loss, offenses, violations of moral imperatives, danger, etc. Because emotions uniquely track these core relational themes by representing them, Prinz satisfies the constraints of uniqueness and intentionality. I’ll explain his theory in detail below.

Prinz argues that emotions are perceptions of bodily changes. By this he means, “they are states within our somatosensory systems that register changes in our bodies” (Prinz, 58). This is very much in line with the Jamesian theory. The suggestion is that emotions uniquely register these specific changes in our bodies. They were “set up” to register such changes. According to Prinz,

Bodily changes occur because they prepare us for response. Our hearts race to increase blood flow, which prepares us for fleeing, fighting, or engaging in other kinds of behavior (Prinz, 68).

We have evolved to be afraid of dangerous things in our ancestor’s environment, for example. Such fear causes us to fight or flight, thus saving our lives. Many emotions evolved because of our cooperation. Anger evolved to be able to punish wrong-doings in society. Guilt evolved to diminish unnecessary anger, and to encourage mercy. Sadness evolved to engender pity and compassion. And the list goes on.
Such responses were either evolutionarily selected or learned, where the learning process depends on a more hard-wired evolutionary one.

But emotions aren’t simply embodied for Prinz (as they are for James). In saying that emotions are perceptions of bodily changes, he means only to say that they are states within our somatosensory systems that register changes in our bodies—they are not yet representations (Prinz, 58). While the Jamesian theory accounts for the embodied part of Prinz’s “embodied appraisals,” we must have representation to account for the appraisal component. According to Prinz, “An appraisal is a representation of the relation between an organism and its environment that bears on well-being” (Prinz, 51). Let’s unpack that a bit.

Prinz adopts Dretske’s theory of mental representation: Mental representations are mental states that satisfy two conditions: (1) they carry information and (2) they can be erroneously applied (Prinz, 53). A state carries information about that with which it reliably cooccurs. Typically, the cooccurrence is causal such that a state carries information about that which reliably causes it. For example, smoke carries information about fire because fire reliably causes smoke. However, carrying information isn’t sufficient for representation; smoke does not represent fires.

To explain how mere information carriers come to be genuine representations, Dretske appeals to the idea that some information carriers have the function of carrying certain information (Prinz, 53). To have the function of carrying certain information, something must be set up to serve that purpose. Compare the
concept of DOG to the smoke caused by fire. The concept of DOG carries the information about actual dogs, just as smoke is caused by fire. However, smoke isn’t set up to be caused by fire—it just is caused by fire. The concept DOG is both caused by actual dogs, and it was acquired—learned—in order to respond to actual dogs. Some mental representations are set up by learning in this way, and others are set up by evolution. For example, our visual system is set up by evolution to respond to edges. And so, a mental representation is a mental state that is reliably caused by something and has been set in place by learning or evolution to detect that thing (Prinz, 54). Prinz’s slogan: “a mental representation is a mental state that has been set up to be set off by something” (Prinz, 54).

To show that emotions are representations, one must show that they are set off by certain things, and that they are set up to be set off by some of those things. As a starting place, Prinz argues that we should accept the premise that emotions are perceptions of bodily changes but deny that they represent those bodily changes (Prinz, 60). Emotions register bodily changes, but they represent something external to us. Emotions represent relational properties that matter to us. For instance, sadness represents loss, where loss is a relational property—it is the elimination of something valued. Anger represents a demeaning offense, where something’s being an offense is a relational property—it diminishes something valued.

Emotions might misfire. Sure, sadness might occur while under the influence of certain drugs, while listening to music, or even while making a sad facial
expression. But this only shows that the proposal can satisfy Dretske’s stricture that representations must be capable of occurring in error (Prinz, 64). But the conclusion remains: emotions represent core relational themes (Prinz, 66). They represent relational properties that matter most to us. This is what they are set off by.

But they were also set up to represent core relational themes—they have the function of representing core relational themes (Prinz, 66). This is because core relational themes are directly relevant to our needs and interests (Prinz, 66). And we have evolved as social creatures to satisfy our needs and interests. Insofar as an emotion serves that purpose, it has the function of tracking the relevant core relational theme.

Furthermore, the representational content of emotions need not be cognitive at all for Prinz. Sadness can occur without the judgment that there has been an irrevocable loss. But sadness represents what that judgment represents—an irrevocable loss. It has the same meaning but a different form (Prinz, 65). How can this be? First, here’s Prinz’s definition of cognition: cognitive states and processes are those that exploit representations that are under the control of an organism rather than under the control of the environment (Prinz, 45). Concepts are mental representations that can come under organismic control. If this definition of concepts is adopted, cognitions can be defined simply as states containing concepts (Prinz, 46). Prinz thinks that emotions are not in fact cognitive, most of the time. They are not generated by acts of cognition, and they are not conceptual (Prinz, 50).
Well, at least they don’t have to be generated by acts of cognition, and they don’t have to be conceptual. Sometimes they are, but this is not an essential feature of emotions.\footnote{It’s not obvious to me that Prinz has argued for this position. I’m not sure I completely agree. I think his arguments show emotions are nonpropositional—they are not judgments—but I’m not sure they show emotions are nonconceptual.}

Appraisal theories claim that emotions necessarily comprise representations of organism-environment relations with respect to well-being (Prinz, 52). An embodied appraisal theory says that such representations can be inextricably bound up with states that are involved in the detection of bodily changes (Prinz, 52). This theory satisfies the uniqueness constraint because emotions are the only states that represent well-being by registering bodily changes. To appraise a situation is to represent it as having some bearing on one’s interests. We can form judgments that appraise a situation—I might judge that my daughter performing well in school satisfies an interest I have. Emotions do not appraise situations in this way. Instead, emotions are unique mental representations that have been set up to be set off by unique stimuli. Prinz considers an analogy with perception, “a state in the visual system \textit{registers} a particular luminance discontinuity, but it \textit{represents} an edge” (Prinz, 58). It represents the edge by registering the luminance discontinuity. And it was set up by evolution to register the luminance discontinuity, but such a system has the function of representing edges. When it comes to emotions, there’s a specific way they feel—bodily. Emotions are set off by such bodily changes. But
they do not have the function of registering bodily changes. They were set up to represent something external to us—core relational themes.

This distinguishes them from judgments—judgments do not represent by registering anything changing in the body. It also distinguishes emotions from other sorts of pro-attitudes such as desires because desires do not represent relational properties. Desires represent states of affairs. The theory also satisfies the intentionality constraint because emotions have content. The content of an emotion is a core relational theme. Such content may be nonpropositional, but it is content nonetheless.

In the following chapter, I explain what it means to have content that is nonpropositional by appealing to Robert Stalnaker’s account of nonconceptual content. If a content is propositional, then it is conceptual (but not the other way around). I think embodied appraisals are nonpropositional (even if not nonconceptual). But I also think emotions as embodied appraisals can be reasons for action. Many philosophers believe that reasons must be propositional in order to play a justificatory role. Here’s my problem: how can nonpropositional emotions play a justificatory role?

In answering this problem I use a parallel solution to a similar worry for perception theorists. Many believe that perceptual experiences are nonconceptual. But perceptual experiences seem to justify perceptual beliefs. How can nonconceptual perceptual experiences play such a justificatory role? If the solution here makes the rational move from nonconceptual content to propositional content,
then it should be even easier for me to make the parallel move from nonpropositional content to propositional content. I show how to make this move in the next chapter.
Works Cited


Chapter 3: An Entitlement Theory

I. Introduction

In the previous chapter, I argued that emotions should be unique, intentional states. Neither strict cognitivism\(^\text{17}\) nor a strict feeling theory satisfied this constraint. Instead, I turned to Jesse Prinz’s theory, an embodied appraisal theory. Such a theory requires both feelings in the body, as well as a representational state, which on this view represents core relational themes. Moreover, I agree with Prinz that the representational states of emotions are nonpropositional. Yet, this dissertation is meant to defend the conclusion that emotions can be reasons for other states as well as for action. How can a nonpropositional state be a reason for anything?

To resolve this, I will draw an analogy between my problem and a problem in perception. Many have wondered how nonconceptual perceptual experiences might justify propositional perceptual beliefs. As an answer to this problem, I will draw on some of Tyler Burge’s work on perceptual entitlement. I will then apply a similar solution to the emotion case. In the same way we’re entitled to take our perceptual experiences

\(^{17}\) Remember, “strict cognitivism” means Solomon’s judgmentalism or Green’s belief/desire pairs.
experiences at face value as reasons for belief\textsuperscript{18}, we are entitled to take our emotional experiences as reasons for beliefs, judgments, and actions. The conclusion is that even if we meet the adequacy constraint with a mental state that is nonpropositional, we can secure the place of emotions as reasons.

II. Stalnaker's Theory of Nonconceptual Content

Before moving on to a solution to my problem, I'd first like to say a bit about what I mean by “nonconceptual content.”\textsuperscript{19} The best view of such content, in my opinion, is from Robert Stalnaker. Stalnaker's theory of content is in the context of perception. There is a wide debate in the perception literature about whether perceptual experiences are constituted by conceptual or nonconceptual content. The original expression of conceptualism comes from Kant's dictum: “thoughts without content are empty, intuitions without concepts are blind.” For the conceptualist, no intentional content qualifies as content at all unless it is structured by concepts that the bearer possesses. I can't have a perceptual experience of scoring a goal in the soccer game without possessing concepts such as GOAL, KICK, BALL, SOCCER, etc. The nonconceptualist disagrees. Many nonconceptualists are impressed by the

\textsuperscript{18} I say my solution to the emotion case is “similar” to Burge's solution to the perception case. They are not exactly the same. Where we part ways is with respect to reasons. Burge argues we are entitled to take perceptual experiences at face value, but the perceptual experience is not a reason to form a perceptual belief. I argue that with respect to emotions, the emotion is a reason to believe, judge, or act.

\textsuperscript{19} Stalnaker's view is about nonconceptual content, but if a content is nonconceptual, then it is surely nonpropositional. So, I think whatever he says here can be directly applied to my problem. My problem of nonpropositional content is in fact even easier than in the nonconceptual case.
phenomenology of perceptual experience, in particular the richness of detail or fineness of grain. Through vision, for example, we are capable of experiencing shades of colors that we do not have concepts for. There is a richness to color experience that we cannot articulate in conceptual terms.

Stalnaker asks: What is the difference between “nonconceptual content” and “conceptual content” as each pertains to perceptual experiences? He claims that all content is an abstract object of some kind. The constituents of one kind of abstract object’s structure might consist wholly of senses or concepts, which are associated with predicates. This would be conceptual content. The constituents of another kind of abstract object’s structure might be individual objects, properties, and relations (the referents of names and the properties and relations expressed by the predicates in the relevant sentences). This would be nonconceptual content. Both kinds of content are intentional. *Intentional content* is a way we are related to the different kinds of abstract objects (i.e., content). Intentional content is neutral as to the kind of content structures because what is essential to intentional content is information. *Informational states* are states in which one thing contains information about another such that there are causal and counterfactual dependencies between the states of one and the states of the other. Stalnaker argues,

To attribute informational content to the state of someone or something is to make a claim about a relation between that person or thing and its environment, and so it is to make a claim that is in part about the environment—about the kinds of things that are found in the environment, and about the way the states of the person or thing are disposed to reflect the properties of those things (Stalnaker, 104).
An object contains information about its environment if the object is in some state that it wouldn’t be in if the environment weren’t a certain way. For example, the pattern of light and dark on the ground on a sunny day carries information about the shape of the tree. This is similar to Dretske’s example of smoke carrying information about fires. Stalnaker claims that such informational states have content (at least in some cases). Presumably they have a kind of content that is in some sense nonconceptual, since it would not be reasonable to attribute conceptual capacities to the patterns of light on the ground. This is what is meant by nonconceptual content: an abstract object that carries information about the environment, where the information is caused by the environment.

It’s important to distinguish between the directional causes of content. Consider an analogy between testimony and the senses. Stalnaker argues that it gets things backwards to explain the content of the witnesses’ testimony in terms of the knowledge state in which it would normally result (in the juror). The witness testifies that she saw the defendant leave the store at 5:30pm. It would be preposterous to explain the witnesses’ testimony by citing the juror’s (subsequent) knowledge that the defendant left the store at 5:30pm. The witness must testify first before the juror can acquire such knowledge. Similarly, it gets things backward to try to explain the way things look in terms of what one would come to believe if one judged that things are as they look. Things must look a certain way before one can judge that things look a certain way.
I think the analogy here can be extended to emotional experiences. Emotions have informational content because there is a relation between them and one’s environment. Moreover, the emotional states of a person are disposed to reflect the relational properties found in one’s environment. Such content may be nonconceptual because it carries information about actual objects, properties, and relations in the environment rather than the concept of such things. Sadness represents loss—a relational property between the subject and the world—it does not represent loss conceptually. Just as it would get things backwards to try to explain the way things look in terms of what one would come to believe if one judged that things are as they look, it would be odd to explain an experience of sadness in terms of what one would come to believe if one judged that there has been a loss. Judging that there has been a loss comes on the waves of sadness, not the other way around. Sure, one may judge that there has been a loss of something valuable, just as one may judge that the tree is 10 feet tall. I can form all sorts of beliefs about core relational themes and bodily changes, just as I can form all sorts of beliefs about my external environment. But the emotional experience “comes first” in the same way a perceptual experience “comes first.”

Why think there is such an analogy between perception and emotion? I think the best argument (for both) comes from young children and nonhuman animals. Just as a small child can have a perceptual experience of a bottle being taken away, so too may he cry when his bottle is taken away without having the concepts of

20 It doesn’t have to be nonconceptual; sometimes it is in fact nonconceptual.
VALUE or LOSS or even BOTTLE. He can experience sadness without judging there has been a loss of something valuable. Dogs can feel ashamed of taking the steak from the table without having the concept of violating an ego-ideal. All sorts of creatures can experience a wide range of emotions without judging anything at all. We can also have a wide range of perceptual experiences without having the appropriate, corresponding concepts. To think otherwise, I believe, is to over-intellectualize both perceptual and emotional experiences.

While I think that Prinz’s embodied appraisal theory best satisfies the two adequacy constraints, and to do so requires that the content of emotions may be nonconceptual, but are at least nonpropositional, I am faced with a larger problem. In chapters 5 and 6, I argue that emotions can be reasons for action. Much debate has been concerned with how a nonconceptual perceptual experience can justify conceptual perceptual beliefs and judgments. I think there’s a parallel problem here: how can a nonpropositional emotional experience justify beliefs and actions? In the next section I turn to Tyler Burge’s entitlement theory of perception to solve this problem for emotions.

III. Burge’s Theory of Entitlement

Tyler Burge offers a broadly externalist account of perceptual belief. Burge views epistemic warrant as a genus with two sub-species. First, there is the epistemically internalist sub-species of epistemic warrant—justification. Justification is warrant that is “conceptually accessible on reflection to the warranted individual” (Burge
That is, justification involves an individual having positive reasons (and being able to access those reasons) for what she believes. Second, there is the epistemically externalist sub-species of epistemic warrant—entitlement. Entitlement is warrant that “need not be fully conceptually accessible, even on reflection, to the warranted individual” (Burge, 2003, 504). That is, a warranted individual need not have positive reasons for what she believes to be entitled to that belief. Burge’s surprising and enticing claim is, then, that “we are entitled to rely, other things being equal, on perception, memory, deductive and inductive reasoning, and on...the word of others” (Burge, 1993, 458). I’ll discuss his entitlement theory, and then I will apply his theory to the case of emotions.

Burge claims that not only are we entitled to rely on perceptual experiences, we are a priori prima facie entitled to rely on them. The entitlement is a priori if its “justificational force is in no way constituted or enhanced by reference to or reliance on the specifics of some range of sense experience or perceptual belief” (Burge, 1993, 458). The entitlement is prima facie in that the entitlement can be defeated or overthrown by counterconsiderations. The counterconsiderations that serve to defeat the original entitlement may be a priori or empirical. That is, it is open to Burge for an empirical consideration to defeat an a priori entitlement. Since we are entitled to rely on perception, we need not be able to articulate or even understand the entitlement to be so entitled. Thus, such an entitlement applies to young children, as well as mature adults.
Burge's entitlement thesis is interesting because Burge believes perceptual experience is *nonconceptual*. Given that perceptual experience is nonconceptual, while perceptual belief is obviously conceptual, the big trick is to transition from perception to belief *in a rational way*. Many have thought perception itself must be conceptual so that it can be a reason for perceptual belief in an inferential, rational way (Sellars, Evans, McDowell). Burge argues that perception cannot be conceptual because some nonhuman animals and young children clearly perceive their environments without possessing the relevant concepts. Surely an infant can see her mother without having the concept MOTHER. Burge argues that his concept of entitlement honors this thought while also giving higher animals and adult humans a rational route from nonpropositional content to propositional content. One is simply entitled to take one's perceptions at face value without believing one's perceptions are veridical or in the absence of a belief that one's perceptual system is reliable. But why are we so entitled?

There are three conditions for entitlement. First, Burge argues, “Epistemic entitlements emerge only where perceptual representations are part of a psychological system that involves beliefs and mental acts” (Burge, 2003, 530). This is rather trivial—beliefs and other propositional attitudes are the kinds of things that are warranted, and so entitlement only applies to such states. Without such states, there would be no entitlement. Second, Burge argues, “To contribute to perceptual entitlement, a perceptual state (type) must be reliably veridical in the perceptual system’s normal environment” (Burge, 2003, 532). A condition on
warrant is that it be a good route to truth (Burge, 2003, 534). And reliably veridical perceptual systems get at how the world really is, that is, what’s true. So, for Burge, any contribution that a perceptual state makes to perceptual entitlement must depend on its having a nature that is reliably veridical (Burge, 2003, 534). Third, perceptual representations have a positive valence insofar as they are presentations of how things are which is beneficial for the animal (Burge, 2003, 542). The positive valence is shown in the fact that perception normally functions to guide animal activity (Burge, 2003, 542). Perception is useful for our survival, which gives some motivation to believe we are entitled to use such representations in higher cognitive functions. So, because perceptual representations are part of a psychological system that involves beliefs, are reliably veridical, and are guides to animal activity, we are entitled to such representations in forming perceptual beliefs.

But there’s more to say about how we rationally transition from nonconceptual perceptual states to propositional beliefs. Simply saying we’re entitled to such beliefs isn’t enough. The transition need not be a temporal transition. They may be caused at the same time. Nonetheless, the perceptual state and the perceptual belief are different psychological states, even if they are caused simultaneously (Burge, 2003, 540). Moreover, the transition is not normally conscious or active. We don’t decide to believe that the grass is green upon perceiving green grass. It just happens—oftentimes without our awareness. And when we are aware, the perception of green grass and the belief that grass is green appears seamless to us. Because there is the possibility of error, the transition is
normative. I might perceive the lines in the Mueller Lyer to be of unequal lengths, but I do not then believe they are (because I measured them, for instance). This would be an instance of a nonveridical perceptual experience and a true belief. We might also have veridical perceptual experiences, and yet false beliefs. I might be in barn-façade land and have a veridical perception of a real barn. Yet, I know I’m in barn-façade land, and so I doubt that it is a real barn—I believe falsely that it is merely a façade. There can be “gaps” between our experiences and our beliefs. And so, since there is a right way and a wrong way to transition from perceptions to beliefs, such transition will be normative.

To preserve the contribution of perceptual states, the transition must meet two further conditions. First, the perceptual belief must be caused by the process that yields the perceptual representation (Burge, 2003, 540). Both the perception and the belief should have a (partly) common causal history, namely the objects or relations in the animal’s environment. Second, the causal relation must not be subject to fluke malfunction, or manipulation from another being.

When the transition goes well, the perception is conceptualized. Through conceptualization the simplest sorts of perceptual beliefs are formed—beliefs that make reference to the same objects, properties, and relations that the perceptual system represents (Burge, 2003, 541-42). We might say that the perception represents the same thing the belief represents, just one is nonconceptual and the other is conceptual.
But there are also ways in which the perception’s contribution to entitlement can be undermined. The perceptual belief might not be causally connected to a relevant perception (Burge, 2003, 543). If you believe there’s a pink elephant running towards you, and there is no elephant there, then your perception does not entitle you to believe there’s a pink elephant running towards you. Or, the perceptual belief might be caused partly by a perception, but might also be manipulated by a further being or be the result of some fluke malfunction (Burge, 2003, 543). This seems to happen when watching a magician. There are veridical perceptions involved, but the magician manipulates your visual system to also make you believe otherwise. Lastly, Burge claims entitlement can be undermined if the transition is hasty, biased or through some misuse on the part of the individual.

None of this is to say an individual is required to believe whatever he perceives. He is entitled to do so, but not required. The individual can have a perceptual representation but withhold belief in the propositional representation normally associated with it (Burge, 2003, 543). An individual might learn not to trust her perceptions under certain circumstances (Burge, 2003, 544). If every time you go to the mall looking to buy some black pants you come home with navy ones, you might start doubting your vision under fluorescent lights. So, you begin to doubt whether the pants are really black, ask for a second opinion, or take them out into the sunlight before purchasing them. Burge thinks such doubt is perfectly acceptable.
I have only one issue with Burge’s entitlement account. In a footnote to “Perceptual Entitlement,” Burge says perceptions are not reasons because reasons must be propositional. The way perceptions entitle us to perceptual beliefs is not the way one belief justifies another—inferentially. I agree the transition is noninferential. But why insist, then, that reasons must be propositional? On a common understanding, R is a reason for P if R causes P, and the agent is warranted in believing P on the basis of R. The same could go for actions: R is a reason to φ if R causes an agent to φ and the agent is warranted in φ-ing on the basis of R. What else could it mean to be a reason? What other condition must be met? The condition is met by perception. A perceptual experience (token) is a reason to believe that P if the perceptual experience causes P (which it does for Burge), and the agent is warranted (entitled) in believing that P on the basis of the perceptual experience.

I think this business about what states we get to call reasons is merely terminological. If Burge wants to stick with reasons as propositional, then I think he’s describing only one sense of the term ‘reason’. The broadest use of ‘reason’ concerns (1) reasons as causes, and (2) reasons as warranting further beliefs, desires, or actions. If warrant can be divided up into justification and entitlement, then there’s no principled reason for thinking ‘reason’ cannot correspond to justification as well as entitlement. So, from now on, I think we can fairly use

---

21 Footnote 27.
'reason’ in a justification for belief or action or as something that entitles one to believe or act. I see no reason for ruling this out.22

So far I have several moving parts at play, and I think it’s time to summarize a bit. I argued in chapter 2 that there is an adequacy constraint on any theory of emotions—emotions must be unique mental states and they must be intentional mental states. Strict cognitivism violates the first condition by reducing emotions to beliefs or desires (or both), whereas the feeling theory violates the second by ignoring the idea that emotions have representational content. Then I introduced Prinz’s embodied appraisal theory as the best theory to satisfy these constraints. But an implication of the embodied appraisal theory is that emotions have nonpropositional representational content. Given that my aim in this dissertation is to show how emotions are reasons to believe and act, the nonpropositional content of emotions is a problem. I need to show how something that’s nonpropositional can both be a reason and also be a reason to believe and act. That’s where Burge’s entitlement theory comes in (after a brief account of what nonconceptual content is in the first place). What I will do now is show that we can co-opt Burge’s solution to the problem for perception to solve the problem I face for emotions.

IV. Emotional Entitlement

Burge argued that “we are entitled to rely, other things being equal, on perception, memory, deductive and inductive reasoning, and on...the word of others” (Burge, 22 Both senses would be distinct from merely explanatory reasons.
1993, 458). I want to add emotions to that list. We are entitled to rely on our emotions in believing and acting\textsuperscript{23}, and when we do, those beliefs and actions are warranted. And if those beliefs and actions are warranted by the emotion as well as caused by the emotion, then we can say that the emotion is the reason for believing or acting.

Again, the entitlement is prima facie in that it can be defeated or overthrown by counterconsiderations. In the case of emotions, such counterconsiderations would be affective disorders, sleep deprivation, brain injury, etc. We do not need to rule out such counterconsiderations in advance. We only need to pay heed once they have come to our attention. Since we are entitled to rely on our emotions, we need not be able to articulate or even understand the entitlement to be so entitled. We need not be aware of the emotion as distinct from a judgment or belief, and the emotion might occur simultaneously with the judgment. Moreover, we can be entitled on an occasion even if a particular emotion has somehow gone awry—so long as our emotional capacities generally are appropriate.

There are three conditions Burge has for perceptual entitlement. First, Burge argues, “Epistemic entitlements emerge only where perceptual representations are part of a psychological system that involves beliefs and mental acts” (Burge, 2003, 530). I think the same goes for emotions. However, it’s not so trivial. While we are entitled to our beliefs and other propositional attitudes on the basis of emotions,

\textsuperscript{23} Throughout this section, I talk about emotions causing and warranting actions. That is really shorthand. Emotions warrant intentions, and then intentions cause actions.
things get tricky for action. You might think that a nonhuman animal or a small child can experience fear, and run as fast as possible in the opposite direction of the danger, without having any propositional attitudes. But there is a question whether such an unsophisticated animal runs out of fear or out of instinct. It is reasonable to think that only when an animal has at least some propositional attitudes can it act for an emotion. If there are no propositional attitudes whatsoever, then the actions are purely caused by instinct. Then, there can be a mix of emotion and instinct when the animal is capable of having propositional attitudes, but no such attitudes are involved in the current action. In either case, being entitled to act arises only when the action is responding to an emotion rather than instinct. If the action is responding to instinct, then the action is merely caused—and not warranted.

Second, Burge argues, “To contribute to perceptual entitlement, a perceptual state (type) must be reliably veridical in the perceptual system’s normal environment” (Burge, 2003, 532). In the emotion literature, such a criteria is fittingness. Fittingness is an external property like Burge’s entitlement. Consider some examples. Fear is fitting just in case the object feared really is dangerous. Sadness is fitting just in case there really was a loss of something valuable. Anger is fitting just in case someone really did offend you. One cannot merely believe the object is dangerous, the loss was real, or the offense occurred. They really had to be there—in the world, so to speak. Fittingness for emotions is akin to truth for beliefs, and both are akin to veridicality for perceptions. They each involve a mind-to-world direction of fit. So it’s no surprise that all of these contribute to knowledge about
the world. Insofar as they contribute to knowledge of the world, they satisfy the epistemic good—a route to truth. Reliably veridical perceptual systems get at how the world really is, and so do reliably fitting emotions. If emotions are reliably fitting, then we are entitled to take them at face value in believing and acting.24

Third, Burge argues that perceptual representations have a positive valence insofar as they are presentations of how things are which is beneficial for the animal (Burge, 2003, 542). The positive valence is shown in the fact that perception normally functions to guide animal activity (Burge, 2003, 542). This is the easiest condition to meet for emotions. Emotions are valenced positively and negatively, and they’ve evolved for the purpose of guiding animal activity. The way emotions feel in our bodies prepares us for action, and those bodily changes are essential to emotional experience. So, because emotional experiences are part of a psychological system that involves beliefs, are reliably fitting, and are guides to animal activity, we are entitled to such emotions in believing and acting.

24 It is important that my conclusion is in the form of a conditional. The antecedent “if emotions are reliably fitting” has not been defended here. It is a classic problem in the epistemology literature to define what counts as a “reliable” perceptual system. Is it all the time? Surely not. Is it 80% veridicality? 75%? More than 50%? There is no answer to this question. The same problem would occur for the fittingness of emotions. Must emotions be fitting all the time? Surely not. I think emotions only have to be fitting more than 50% of the time to count as reliable. Do emotions satisfy that requirement? Well, one might tell a story about the valence of emotions, which amounts to pleasure and displeasure. And drawing on Tim Schroeder’s Three Faces of Desire, we might say that pleasure and displeasure reliably represent how we’re faring in the world. This could be the source of the reliability of emotions. However, defending the antecedent is not necessary for my argument here. Moreover, attempting to defend it here would take me too far away from my central argument; it could be a dissertation in and of itself.
Now I need to say how we transition from nonpropositional emotions to propositional beliefs, as well as actions. When considering judgment alone, just as a perceptual state and a perceptual belief are different psychological states even if they are caused simultaneously, so too might fear and a judgment that something is dangerous be caused simultaneously. But that does not entail that the emotion is the same thing as a judgment. The relation between emotion and action is different. Some emotion theorists (Affect Program Theorists) include action tendencies in the nature of emotion itself. In that case, the emotion and the action might be caused simultaneously. But on the embodied appraisal theory I advocate here, relational properties cause emotions, and emotions cause actions (or judgments). But notice that actions are not propositional states, so the question of entitlement will be different. I’m not concerned, however, with any action whatsoever. I’m concerned with intentional action. Intentional actions are done with intention, where intentions are propositional attitudes.\textsuperscript{25} Since intentional actions involve propositional attitudes necessarily, then we can direct our attention to how we transition from nonpropositional emotions to propositional intentions (the backbone of intentional action).

Again, because there is the possibility of error, the transition from emotion to belief or intention is normative. There are many examples of fitting emotions causing true beliefs in the right sort of way. When I grow jealous of my partner’s

\textsuperscript{25} I’m not concerned here with whether intentional actions are done for intentions, as a response to intentions, or are caused by intentions. I don’t think it matters much.
suspicious activity with his secretary, when my emotion is fitting (he is having an affair with her), that can cause me to believe truly that he is cheating on me. But we also have instances of fitting emotions causing false beliefs, as when someone has experienced a brain injury that affects his emotional responses. Such a person doesn’t generally trust his emotions, but on an occasion his anger is a fitting response to an offense. Yet, because of his mistrust, he fails to judge that someone has offended him, and he fails to intend to seek revenge. And so, since there is a right way and a wrong way to transition from emotions to beliefs and intentions, such transition will also be normative.

The analogy continues. To preserve the contribution of perceptual states, the transition to perceptual belief had to meet two further conditions. First, the perceptual belief must be caused by the process that yields the perceptual representation (Burge, 2003, 540). Both the perception and the belief should have a (partly) common causal history, namely the objects or relations in the animal’s environment. The same goes for emotions. Just as Prinz notes, emotions and judgments represent the same thing—the relational properties of objects—but emotions represent them nonpropositionally while judgments represent them in the form of a proposition. But the same thing—the object, relation, or event in the world—causes both the emotion and the judgment. Second, the causal relation must not be subject to fluke malfunction, or manipulation from another being.

Just as entitlement to perceptual beliefs can be undermined, so too can beliefs and actions based on emotional experiences. The belief or action might not
be causally connected to a relevant emotion. If you believe there is a dangerous snake on the ground, and run from it, but it’s really a water hose coiled up on the ground, then the causal connection is faulty, and your entitlement is undermined. Your entitlement can also be undermined by known defeaters—affective disorders and the like.

None of this is to say an individual is required to believe or intend whatever he experiences emotionally. He is entitled to do so, but he doesn’t have to. An agent might learn to mistrust her emotional responses to her father (“Daddy issues”), while trusting those same emotional responses to others and in different contexts. Again, this doesn’t eliminate all entitlement; it is a defeater for particular instances of emotions.

The conclusion I draw is this: emotions cause beliefs and intentions, and they warrant those beliefs and intentions. We are entitled to believe and intend on the basis of our emotional experiences. This is all that is required to call something a reason. And so, just as we can have perceptual experiences as reasons to believe, we can say emotions are reasons to believe and intend.

V. Concluding Remarks

I hope that the idea that emotions can warrant beliefs and intentions has lost some of its initial sting. Many people would deny such a claim. But, I do not take myself to have fully defended that claim here. Instead, I’ve merely set up the prerequisites for such a defense to come. In order for emotions to be reasons for belief and intention,
they must have certain properties that entitle us to rely on them. Moreover, when I use ‘emotion’ throughout this dissertation, I am speaking of the embodied appraisal theory from Prinz. I think it’s necessary to make this clear at the outset—and I’ve tried to argue for it as the best option.

The remainder of this dissertation defends the thesis that emotions are reasons for beliefs and intentions. Although emotions warrant beliefs and intentions noninferentially, they are nonetheless reasons for those further states. The normative move from emotions to beliefs and intentions is that of entitlement. This view that emotions are reasons challenges the standard account of intentional action: an account on which only belief/desire pairs can be reasons for action. In the next chapter, I will present the standard account, raise a challenge to it from actions expressing emotions, and show how attempts to save the standard account fail. Then in subsequent chapters I will argue for my own positive view of emotions rationalizing actions.
Works Cited


Chapter 4: On Actions Expressing Emotions

I. Introduction

There’s a predominant view in metaethics and philosophy of action that I believe is inadequate. I shall call this view the “standard account” of intentional action. There are a few common core claims of the standard account. The standard account claims that intentional actions are actions done for a reason. Reasons are the kinds of things that rationalize actions. At the minimum, reasons rationalize action in the sense that from the agent’s point of view, there was something to be said for acting. This is the so-called “intelligibility requirement.” Reasons have to be a certain way, metaphysically, to be able to play this rationalizing role. First, the reason must, in part, include a motivating state (usually desire) to explain why the agent acted. This is because intentional actions are (said to be) necessarily goal-directed, and having a goal just is being motivated. Second, the reason must, in part, include a belief about the means for obtaining the thing being pursued by the agent. The belief may be false. Moreover, actions are rationalized only if the agent acts for a reason, or

26 I’m trying to capture here what is in common between Davidson, Williams, Smith, and Arpaly.
because of the reason, or by virtue of the content of the reason. The action must be caused by the reason in the right sort of way so as to make the agent seem rational or reasonable. This “right sort of way” is often understood in an inferential way. In a nutshell, here’s the standard account:

**Standard Account:** For all actions, A’s action is intentional if and only if she does it 

\[ \text{for a reason where performing the action for a reason requires (1) A has some element in her subjective motivational set (usually desire) the satisfaction of which will be served by her } \varphi \text{-ing, (2) A has a belief about the means for satisfying her desire, and (3) A’s action involves a practical inference and not mere content-efficacy.} \]

The purpose of this chapter is to raise a challenge to the standard account from expressive actions. Expressive actions are those that express some occurrent emotional state, which qualify as intentional actions, and yet they do not seem to meet the criteria set forth by the standard account. Examples include hiding under your sheets out of fear, patting your child on the head out of love, jumping up and down out of joy, or kicking the table out of anger.

I begin by clarifying the challenge to the standard account from expressive actions. This challenge comes primarily from Rosalind Hursthouse. She argues that expressive actions are intentional, and yet not done for reasons, as the standard account understands “done for a reason.” I then present some plausible solutions to the challenge. Some have argued that expressive actions are not intentional, and so they do not undermine the standard account. Others have tried to merely tack on

---

27 More on the difference between practical inference and content-efficacy will follow.
emotions to the standard the belief/desire pairs. Some attempt to substitute some other psychological state for beliefs, namely imaginings. The main purpose of this chapter is to object to each of these accounts, showing how they do not adequately answer Hursthouse’s challenge. I conclude that the standard account, as stated, cannot be saved from the challenge of expressive actions.

II. The Challenge

What are expressive actions? Well, the class of expressive actions is wide and diverse. They all have the expression of an emotion in common, but such expressions might range from smiling and laughing from happiness, to clenching your fists in anger, to hiding under the bed sheets in fear, to covering your face in shame, to caressing the clothes of your deceased spouse in grief. According to Peter Goldie, these expressive actions lie between bodily changes which are part of an emotion and reasoned actions out of an emotion (Goldie, 25). At one extreme are bodily changes, which are autonomic nervous system responses and hormonal changes such as “sweating, change of heart-rate, secretion of adrenalin and so forth, and muscular reactions such as trembling, flinching, and so forth” (Goldie, 25). These are changes that just happen to us. We can’t make our heart race or palms sweat. These are not actions at all. On the other end of the spectrum are reasoned actions out of an emotion. By reasoned actions, Goldie means “those actions which can be adequately explained or rationalized by reference to appropriate combinations of beliefs and desires” (Goldie, 25). For example, a tiger is stalking
you in the wild, and you jump over a fence. Your act of jumping over the fence can be explained by your desire to get away from it combined with your belief that you can best get away from it by jumping the fence.\textsuperscript{28} Goldie thinks in such cases the fear isn’t even necessary to explain the action. \textit{Expressive actions} are in the middle between bodily changes and reasoned actions. Expressive actions are more than bodily changes because they don’t merely happen to us, and yet they are less than reasoned actions because they are somewhat less than fully rationalizable in terms of combinations of beliefs and desires. The challenge is to show how expressive actions are still full-fledged intentional actions, and yet not rationalized by beliefs and desires.

The classic example of an expressive action comes from Hursthouse. She asks us to consider “Jane, who, in a wave of hatred for Joan, tears at Joan’s photo with her nails, and gouges holes in her eyes” (Hursthouse, 59). Suppose Jane and Joan are best friends, and Jane has just discovered Joan has been sleeping with her husband. Jane’s rage is growing as she scans her surroundings, and sees a picture on the mantle of happy Joan with her and her husband. She flies across the room, grabs the picture, and violently scratches out Joan’s eyes. What should we make of Jane’s action?

I hope it seems obvious that what Jane does is more than mere bodily changes. No mere autonomic systems are flailing her across the room. It seems that

\textsuperscript{28} This is Goldie’s explanation. I will disagree with this explanation later. I believe the emotion plus a desire can explain the action.
what Jane does is an action. She sets her focus on the picture of Joan—the object of her rage—and goes after it. Hursthouse and Goldie agree that we can go so far as to say her action was intentional. What’s the difference between Hursthouse and Goldie? How is Jane’s action a challenge to the standard account? Well, in Hursthouse’s terminology, “According to the standard account of actions and their explanations, intentional actions are actions done because the agent has a certain desire/belief pair that explains the action by rationalizing it” (Hursthouse, 57). The problem with Jane’s action is that her particular beliefs and desires do not explain her action by rationalizing it. So, we have an intentional action, but a belief-desire pair does not rationalize the action. That means not all intentional actions are rationalized in the way the standard account says. And so, we seem to have a counterexample to the standard account. The standard account says intentional actions are explained by beliefs and desires which rationalize the action. Expressive actions are intentional but NOT explained solely by beliefs and desires. So, according to Hursthouse, expressive actions challenge the standard account of intentional action.

II. a. Raz’s Solution

Joseph Raz, however, challenges the obviousness of Jane’s action being intentional. For Raz, if an expressive action is intentional, then the action must be under the control of the agent (Raz, 40). The initiation of the action should be up to the agent, and the continuation of the action must be under the control of the agent. For
example, suppose Jane chooses a photograph that is unframed rather than one behind glass. Such a choice would show she is in control of the initiation. While scratching out the eyes, she is careful not to also scratch her antique dining room table (suppose she intentionally places the photo on a thick placemat). Such choices show some degree of control over the action. So, for expressive actions to be intentional, the action must be under the control of the agent.

But Raz argues that expressive actions involve a loss of control, and this loss makes them less than fully intentional. Jane probably used much more force in scratching out the eyes than she wanted to (with the unintended consequence of ruining the placemat). Raz thinks expressive emotions “explode from within without [any] element of calculation” (Raz, 42). Hand-in-hand with this loss of control is a loss of intentionality. In Raz’s words: “The crucial point is that the very loss of control which makes them cases of letting oneself act for a reason rather than cases of acting for a reason also makes them somewhat less than typically intentional” (Raz, 44). So long as there is a correspondence between the loss of control and loss of intentional action, and expressive actions involve a loss of control, then expressive actions are less than fully intentional. Raz’s response to Hursthouse, then, is that expressive actions necessarily involve some loss of control, and so they are not fully intentional. If they’re not fully intentional, then we do not have a class of intentional action that’s not done for a reason.

---

29 It’s not so clear what “letting oneself act for a reason” amounts to. It’s supposed to contrast with acting for a reason, where acting for a reason involves the will, while letting oneself acting for a reason does not.
But Hursthouse’s challenge was a class of actions that are intentional and not done for a reason at all. At most, Raz’s arguments show expressive actions are a little less than fully intentional, and so the agent is a little less than fully acting for reasons.\textsuperscript{30} As I’ve construed the standard account, expressive actions are supposed not to be done for a reason at all, and so Raz’s position with respect to reasons is already at odds with it. And so long as expressive actions are intentional at all—more intentional than mere behaviors—then we have a class of intentional action that is not done for a reason, understood in the standard belief/desire way. So, I don’t think Raz’s argument here is enough to show expressive actions are nonintentional, or mere behaviors. It is interesting to think about what sort of loss of control there is in some expressive actions, but such a loss of control is insufficient to show there are no intentional actions that are not done for a reason.

\textit{II. b. The Belief/desire Solution}

But consider another way that a proponent of the standard account might reply to Hursthouse’s challenge: Surely Jane has a desire to gouge out the eyes in the photograph, and this combines with a belief that she can do this by doing this-and-that with her fingernails. Isn’t that enough to place her action squarely within the boundaries of the standard account of action? This would seem to be an appropriate explanation of action according to the standard account. In fact, it’s the only

\textsuperscript{30} This implies that acting for a reason is a gradient concept. I think what Raz means here is the difference between “acting for a reason” and “letting oneself act for reasons” where the latter is somehow less than the former.
explanation of Jane’s action that would be consistent with the standard account. Presumably, Jane does have a desire to damage the photograph of Joan or else she wouldn’t be directing her attention and efforts at it. And she has a variety of beliefs about how to go about damaging it. Maybe she believes scissors are the most effective instruments, and so that’s why she’s using them rather than fingernails or a knife.

But I think many of us would be unsatisfied with the foregoing explanation. Suppose I walk into the room and see Jane scratching up those photographs. I ask her why she’s doing that, and she simply says “because I want to scratch out Joan’s eyes in the photograph, and using the scissors in this way is the best way to do that.” I wouldn’t feel as if she’d explained her action at all. No—she’d just be telling me what she’s doing. She wouldn’t be explaining why she’s doing it, which was what I wanted to know. Something’s missing.

These particular beliefs and desires do not combine to be a reason why Jane is doing what she’s doing. They do not work to rationalize Jane’s action, in the sense that merely citing her belief and desire do not make her action any more intelligible. Rather, what we might need is some reference to her emotion. What’s missing is the way in which Jane is scratching up those photographs—angrily. Jane is expressing her rage in an intentional action. But that intentional action cannot be adequately explained by only reference to beliefs and desires.

Michael Smith agrees that this belief-desire pair fails to rationalize Jane’s action, and so he supplements the belief and desire with an emotion. Smith
considers another one of Hursthouse’s examples wherein a grieving man is found rolling around in his recently deceased wife’s clothes. Smith grants that there’s a simple belief/desire pair explanation of this action: “the man is doing what he is doing because he desires to roll around in his dead wife’s clothes and believes that he can do so by doing just what he is doing: that is, by rolling around in those particular clothes that he is rolling around in” (Smith, 22). But, as Smith says, this “Humean explanation is distinctly unsatisfying” (Smith, 22). It is dissatisfying because it immediately prompts the further question: “And why would anyone want to roll around in his dead wife’s clothes?” (Smith, 22). The answer Smith gives is “because he is grieving for her” (Smith, italics mine, 22). The grieving man’s emotion—his grief—is supposed to make sense of his belief/desire pair. We explain the grieving man’s beliefs and desires by citing his grief. And according to Smith, “the explanation explains because it takes the fact that there is a Humean explanation for granted” (Smith, 22). That is, a Humean of the sort that accepts a belief/desire model of action (what I’ve been calling the “standard account”) can supplement his view with emotions so long as the explanation in terms of emotion takes for granted some Humean explanation. I think the way in which the emotion explanation takes a Humean explanation for granted is that the emotion makes sense of the belief/desire pair rather than making sense of the action itself. It’s once removed from the action. However, the belief/desire pair makes sense of the action itself. So, the emotion explains the belief/desire pair, but only belief/desire pairs

31 Throughout, I take this example to have a parallel structure to the Jane example.
can explain action. The man’s grief explains why he desires to roll around in his dead wife’s clothes and believes that he can do so by doing just what he is doing: that is, by rolling around in those particular clothes that he is rolling around in. But this desire and belief combine together to explain his action: rolling around in his dead wife’s clothes. So, we still maintain a standard account—the action is explained by beliefs and desires. It’s just that those beliefs and desires are further explained by an emotion.

According to Goldie, there’s a problem, however, when we try to apply Smith’s solution to the case of Jane. He argues, “it does not really make it at all clear why someone should be disposed when hating or in anger to do such a “bizarre” thing as scratch the eyes in a photo of the person they hate or are angry with” (Goldie, 28). The reason why the belief-desire rationalization was inadequate was because it failed to explain why Jane would be doing what she’s doing. Goldie’s criticism of Smith is that the emotion equally fails to explain why Jane is desiring and believing those things. We’ve simply moved the problem back one stage—from a failure to explain the action to a failure to explain the belief/desire pair. There’s no obvious connection between being angry with someone and scratching out their eyes in a photograph. Anger does no better than beliefs and desires. So,

---

32 I’ll return to Smith’s solution in Chapter 6.
33 I will question this as actually a standard account in the next chapter.
34 The same problem arises for the grief situation. The desire to roll around in one’s dead wife’s clothes combined with the belief that these are one’s dead wife’s clothes doesn’t really explain why one would be rolling around in such a way.
we need to look elsewhere for the right kind of connection. Goldie looks to the imagination.

II. c. Goldie’s Solution

Goldie essentially broadens the standard account of action to say, “an action can, quite generally, be explained and understood as rational under a description by the attribution to the agent of a combination of some sort of pro-attitude, such as desire, and some sort of cognitive state, such as belief” (Goldie, italics mine, 26). Such a rationalizing explanation gives the reason why the agent acted, but the states doing the rationalizing need not be restricted to only beliefs and desires.

The point of Goldie’s paper is to refute Hursthouse’s challenge: we do not have a class of intentional action that cannot be explained by the standard account, properly understood. Hursthouse understood the standard account as too narrow—only focused on beliefs and desires. Goldie thinks we can broaden our understanding of the standard account to include the combination of cognitions and attitudes of other sorts. In particular, considering expressive actions leads us to broaden the standard account to include imaginings as “belief-like” and wishes as “desire-like”.

For typical belief-desire explanations, you have an agent who wants something—a sandwich, say, and then the agent forms specific beliefs about how to go about getting the thing she wants. We’ve seen already how such explanations are inadequate for expressive actions, though. Instead, Goldie argues that expressive
actions are best explained by a wish. For Goldie, “when I wish for something, I desire that thing, and I also imagine, or am disposed to imagine, the desire to be satisfied” (Goldie, 28). A wish combines the pro-attitude of a desire with the cognitive state of imagining that desire to be satisfied. We can then explain Jane’s action: “Jane has a desire to scratch out Joan’s eyes, and the imagining is that she is doing this in the expressive action” (Goldie, 29). That is, as Jane scratches out the eyes in the photo, she is wishing she were scratching out Joan’s actual eyes (i.e., she’s imagining she’s acting on her actual desire).

Goldie, however, must answer the challenge he raised for Smith: what explains such a bizarre action as scratching out the eyes in a photograph? Goldie argues that in situations like Jane’s we have the choice between attributing to Jane the desire to scratch out Joan’s eyes in the photograph and the desire to scratch out Joan’s actual eyes. Goldie says that scratching out one’s enemy’s actual eyes is primitively intelligible, whereas scratching out one’s enemy’s eyes in a photograph is not. By ‘primitively intelligible’, Goldie means something like the typical desire response from certain emotions. For example, desiring to harm someone out of anger is primitively intelligible. We all understand it without further explanation. Or desiring to be away from the object of fear is primitively intelligible. Or desiring to hide one’s face in shame is primitively intelligible. When we attribute to Jane the desire to scratch out Joan’s actual eyes, “we have a real psychological state, one which is primitively intelligible for Jane to have towards her rival whom she hates and has feelings of anger towards” (Goldie, 29). And there’s an underlying
assumption here that it is better to attribute primitively intelligible desires to people if at all possible. So, we should say Jane has a desire to scratch out Joan's actual eyes, and she is imagining she is doing this as she scratches out the eyes in the photograph. That is, she is wishing she were scratching out Joan's actual eyes.

Goldie's idea here solves two problems. First, it is pretty bizarre to just have a desire to scratch the eyes out of one's rival in a photograph. We might understand when people do it. But there is still something bizarre in having just this desire without some deeper desire to harm the actual person. This was the problem with the standard account explanation as well as Smith's solution—no one could explain the desire to scratch the eyes out in the photograph. Goldie thinks he's got such an explanation, and this is a merit of his view. He supplies a deeper desire for Jane—a desire to scratch out Jane's actual eyes. And this desire is supposed to explain her expressive action via its connection to the imagination. Second, it would be weird if we straightforwardly combined Jane's desire to scratch out Joan's actual eyes with a standard account style means-end belief that by scratching out the eyes in the photograph, Jane is satisfying that desire. We need not attribute such a weird belief to Jane—we should not be put in a position to say Jane believes that by scratching out the eyes in the photograph of Joan, she is thereby scratching out Joan's actual eyes. That would be crazy! Goldie need not say such crazy things. Instead, he says that the desire to scratch out Joan's actual eyes combines with an imagining that Joan's actual eyes are being scratched out rather than with an actual belief Joan's eyes are being scratched out. That's far less crazy. Furthermore, the imagining is
then part of the rationalization of the action, and so Goldie’s view is just a slightly broader version of the standard account.

So, Goldie explains expressive actions in terms of a wish, and this explanation is supposed to save the (slightly modified) standard account of intentional action from Hursthouse’s challenge. Goldie is careful when he states the standard account. He says action can be “explained and understood as rational under a description by the attribution to the agent of a combination of some sort of pro-attitude, such as desire, and some sort of cognitive state, such as belief” (Goldie, italics mine, 26). He leaves it an open question whether other sorts of pro-attitudes and cognitive states might combine together to rationalize action. And that’s exactly what he does with a wish. A wish is the combination of a pro-attitude, desire, and a cognitive state, imagining. We want something, and then imagine it to be the case when we wish. And then, in certain circumstances, this imagining-desire combination rationalizes action. Thus, expressive actions—explained in terms of a wish—are perfectly consistent with the standard account of intentional action.

III. Objections to Goldie’s Solution

In this section, I’ll raise a series of problems with Goldie’s account. First, I’ll argue that Goldie is wrong about what wishes are, conceptually, and that if we properly understand the concept, we can really do without it. Second, I consider a better version of the imagining/wish explanation of expressive actions that comes from Velleman. But even this better account faces a crucial problem: not all expressive
actions can be explained in terms of imaginings and wishes. So, it seems the challenge to the broader standard account still stands. Lastly, I argue for the further claim that even those expressive actions that can be explained by imaginings/wishes cannot be rationalized by such mental states. So, the imagining/wish account cannot save the standard account of action from Hursthouse’s challenge. I’ll then go on to argue that what’s missing from the imagining/wishes account is the same thing that’s missing from the belief/desire account: reference to an emotion.

III. a. Goldie’s Account of a Wish is Unsatisfactory

Goldie says a wish is a state that combines an imagining with a desire. However, I think we ought to maintain a firm distinction between wishing and desiring. So, I think it’s a mistake to say wishes have desires as components. Here is the main distinction: desires guide action while wishes do not. When I wish I had a new car, I sit around idly dreaming of a new car. When I desire a new car, I research cars online, go to used car lots, talk to friends about the reliability of their cars, etc. My desire guides my action; my wish does not. Sure, you might say you wished for a new car as you troll the used car lot. But if you are actively looking for a new car, then you’re desiring one, not wishing for one. When Goldie says that wishes are composed of desires, he has not honored this distinction. Moreover, Goldie wants his concept of a wish to be action guiding. So, as I shall argue below, he ought to drop talk of wishes and keep only talk of desires.

35 Thanks to Abe Roth for help with this distinction.
Goldie wants to expand the standard account to include some sort of cognitive state, such as belief, and some sort of pro-attitude, such as desire. That expansion requires two states. But when he explains expressive actions, he simply explains them in terms of a wish. The wish state is like a sort of “short hand” for imagining + desire. Well, which state is doing the explanation? Do wishes explain expressive actions? Or do imaginings combined with desires explain them? I don’t see why Goldie has to name the combination of imagining + desire a “wish” when it seems like all the work is being done by the imagining/desire pair, in the same way non-expressive actions are explained by belief/desire pairs. So, Goldie has gone out of his way to give a truncated definition of wishes. He gets it wrong what a wish is, and he doesn’t even need to discuss wishes at all. He could get everything he wants by simply talking about imagining/desire pairs. Henceforth, I will refer to Goldie’s explanation as an “imagining/desire pair” rather than as a “wish.” I think this is more closely what Goldie means, and it is less ambiguous with respect to the term “wish.”

III. b. Not All Expressive Actions are Explained by Imagining/Desire Pairs

So we grant Goldie some sort of state, call it an ‘imagining/desire pair’. Does an imagining/desire pair capture the nature of expressive actions? Consider Jane again. We’re told by Goldie that Jane has a desire to scratch the actual Joan’s eyes out, and we’re told that Jane is imagining she is doing this while she scratches the eyes out of Joan in the photo. I have two questions for Goldie in this section: (1) why
should we think Jane has this desire? (2) Why should we think Jane is imagining anything? I’ll take each of these questions in turn.

As for question (1), why should we think Jane has a desire to harm the actual Joan? Is this a requirement for all expressive actions? Consider an alternative explanation of Jane’s action. Jane discovers a love letter planning a romantic vacation between her husband and Joan. After she crumbles the note, she frantically scans the room. There—on the mantle—Jane sees the photo of Joan, herself, and her husband. The mere sight of a happy smiling Joan rips at her core. She charges at the photo, and wants to take the happy look out of those eyes. Jane might want simply not to see that face right now. She might want the picture of Joan to be mangled like she feels torn apart now. She might desire to destroy something, in rage, and the picture of Joan just seemed to be the obvious choice to her at the time. She might go on to destroy a pillow, too. Who knows? There are many candidate desires Jane has that would contribute to an explanation of her action. The desire to harm the actual Joan is mere supposition.

According to David Velleman, Jane might imagine she’s scratching out Joan’s eyes without having a desire to do this at all. She might be pretending she’s scratching out Joan’s eyes without wanting to do so. If Velleman’s theory works out, then the result is an imagining/desire pair that explains Jane’s action without having to attribute to Jane the desire to harm the actual Joan.

36 The same question could be asked in the grief case: why think the grieving man wants to be rolling around with his actual wife?
Velleman offers an account of pretending in “On the Aim of Belief” that might help Goldie out. Rather than thinking of Jane as desiring to scratch out the actual Joan’s eyes as she tears at the photograph, it would be better to think she’s pretending she’s doing this. Consider the following case. Velleman asks us to think of pretending to be an elephant by swinging your arm in front of your face as if it’s a trunk. Suppose you then walk over to a chair and “drink” out of it as if it’s a pail of water. How ought we explain this behavior? Velleman argues we should think of your point-of-view from within the fiction. Obviously you know that the object in front of you is a chair, and not a pail of water. But you imagine the chair is a pail of water, and—from inside the pretend elephant’s mind—you believe it is a pail of water. You have a mock-belief that this is a pail of water. And this mock-belief explains why you drink from the pail of water as an elephant would.

We must distinguish between the two subjects: “I” and “I-as-an-elephant.” When “I” am pretending to be an elephant, I am imagining I have a trunk and that the chair is a pail of water. When I swing my arm beside the chair, I am pretending to get a drink of water. But when the subject is “me-as-an-elephant,” from within the fiction, then when I swing my trunk by the pail of water, it’s because I desire a drink and believe that I can get a drink from this pail. These are not real beliefs and desires, of course. They are the pretend beliefs and desires of the pretend elephant. So, they are mock-beliefs and mock-desires. I might desire to imagine I am an elephant, and this real desire would explain why I’m holding my arm in a way to

37 “Pretending” in Velleman’s sense to be described.
look like an elephant’s trunk over a chair imagined to be a pail of water. But within
the fiction—from the point-of-view of the elephant I’m pretending to be—I might
desire to get a drink of water. And within the fiction, this mock-desire explains why
I’m holding my trunk this way above the pail of water. So, for Velleman, we have to
distinguish the beliefs and desire of the agent doing the pretending from the mock-
believes and mock-desires that arise out of the fiction itself.

How might Velleman’s account of pretense help here? Consider Jane again.
Suppose Jane is pretending she’s scratching out Joan’s eyes as she’s scratching out
the eyes in the photograph. If this is a real case of pretending, then we have two
subjects: Jane and Jane-who-is-pretending. Jane has an actual desire to pretend
she’s scratching out Joan’s actual eyes while she’s scratching up the photo. And
Jane-who-is-pretending has a mock-desire to scratch out Jane’s eyes, and this mock-
desire combines with a mock-belief that these are Jane’s real eyes to motivate her
behavior. Jane has a desire to pretend to do something, and within the fiction, the
Jane-who-is-pretending is scratching out Joan’s eyes. Nothing here requires we
attribute to Jane the desire to actually harm the real Joan. The actual Jane wants to
pretend to harm Joan, and the pretending Jane only has a mock-desire to harm the
actual Joan. So, we get an explanation of Jane’s action in terms of the imagination,
and yet we need not attribute to Jane a desire to harm the actual Joan. This seems
like progress.
So, Velleman has given us a way to think of Jane as imagining she’s harming the actual Joan without a desire to harm the actual Joan. Maybe that answers my first question: it is not necessary for Jane to have a desire to harm Joan.

As for question (2), why must we include the imagination at all in the explanation of Jane’s action? We need not accept that she must be imagining she is harming the actual Joan when she scratches the photo. Jane might be focused entirely on the photo of Joan, and attempting just to scratch out the eyes in the photo, without imagining scratching the real Joan’s eyes at all. She might be imagining her husband wrapping his arms around Joan lovingly. Or she might be thinking of the two of them kissing or doing loving things. She might not be imagining anything at all. There’s no psychological pressure on Jane to imagine she’s scratching out the actual Joan’s eyes as she does it to the photograph.

I acknowledge that Goldie’s explanation of Jane, and Velleman’s similar explanation, is one possible explanation amongst many. It is possible that Jane deeply desires to scratch out the actual Joan’s eyes, and she is imagining she is doing this while she scratches out the eyes in the photograph. Likewise, it’s possible that Jane is pretending she’s scratching out Joan’s eyes, but this only involves mock-desires, and so Jane has no real desire to harm the actual Joan. However, it’s not necessary to posit the imagination to explain all expressive actions (or even to explain Jane’s expressive action). The Goldie/Velleman explanation simply won’t generalize to other expressive actions. It’s not the best explanation for a wide variety of cases.
Suppose Jane instead expressed her anger by screaming as loud as she can. This isn’t a destructive kind of expression, and so it’s difficult to see how a desire to harm Joan combined with imagining the screaming is harming her would ever be the right explanation. But also consider expressions of other emotions. What about shame? Suppose John is ashamed of being fat at the pool.\textsuperscript{38} Out of this shame, John runs to the bathroom and covers his face—even when no one can see his face in the first place. In that bathroom stall, he’s not imagining someone can see his face, and then covering it. No, he simply covers his face out of shame. What about love? When a father scoops up his young daughter and tosses her up in the air, he’s not imagining she is happy while desiring her happiness. He is simply acting out of love. In each of these cases, and countless others, there is no designated object that is \textit{symbolic} of the real object of the emotion. The person is simply acting out of the emotion without imagining anything.

And so, it seems Goldie’s explanation in terms of an imagining/desire pair, as well as Velleman’s explanation in terms of mock-beliefs and mock-desires, only accounts for a small subset of expressive actions. Their views (so far) only account for those expressive actions with a \textit{symbolic} character.\textsuperscript{39} Goldie claims that, “there is often some symbolic match or correspondence between the object of the emotion and the object towards which the expression of that emotion is directed” (Goldie, 30). Another example of symbolic expressive action would be the man who rolls

\textsuperscript{38} A lovely example from Justin D’Arms.
\textsuperscript{39} And not even all of those.
around in his dead wife’s clothes out of grief. The clothes are symbolic of his late wife. Another example would be when we assign an object the identity of the person we’re angry with—say this cushion or that vase is his face—as we smash or hit it. You imagine that the pillow is your nemesis as you punch it repeatedly. But this isn’t always the case. Sometimes we just kick the table out of anger because it’s nearby and fairly durable, and not because it’s a symbol of whomever it is that we’re really angry with. Other times we jump for joy and dance around the room out of happiness, and there is nothing symbolic about it. I’d go so far as to say most of our expressions of emotions are not symbolic in nature. So, an imagining/desire pair has nothing to say about them.

So, we have two types of expressive actions. One type is symbolic, and thus is explained, on occasion, in terms of cognitive imaginings and conative desires. I’ll grant to Goldie that some expressive actions might, on occasion, be explained in this way. But the other common type of expressive action is not symbolic, and it’s not explained in terms of an imagining/desire pair. And this class of expressive actions continues to pose a major challenge to the standard account of intentional action.

**III. c. Non-symbolic Expressive Actions**

Velleman, however, does consider some expressive actions that are not symbolic. So, maybe he will be helpful again. He gives an example (taken from Hume) of a person who is suspended very high up in a secure cage. A person in such a scenario might tremble in fear despite knowing he is completely safe, i.e., there is no danger
of falling. Velleman’s point is that “although the person doesn’t believe that he’s going to fall, he does imagine falling, and imagination can arouse the same emotions as a belief” (Velleman, 270). Velleman continues the example. Suppose when the man is lowered to the ground, he rattles the bars of his cage in his impatience to get out. The man probably does not believe that he can get out faster by rattling the bars. Instead, his rattling the bars is explained by “the wish that he could escape from the cage more quickly, and he will be imagining a quicker way out” (Velleman, 270). The point generalizes for Velleman: all expressive actions are motivated by wish and imagination, rather than desire and belief. This goes for symbolic expressive actions as well as non-symbolic ones. Of symbolic expressive actions, as well as others such as scratching your head when you're puzzled, or hiding your face in shame, or clenching your fists in anger, or yelling at referees in a sporting event, Velleman argues:

“I know of no satisfactory explanation of these behaviors in terms of desires and beliefs. I can of course concoct desire-belief explanations for them: desire-belief explanations are all too easy to concoct. But the resulting explanations aren’t satisfactory, because they make your behavior look realistically purposive, when it is in fact utterly fantastic” (Velleman, 271).

So, Velleman thinks that non-symbolic expressive actions are explained in terms of imagining + wishing whereas it seemed Goldie couldn't say that. When you yell at the referee, you don’t believe your expression of anger will affect anything, but you imagine and wish it would. When you roll around in your dead wife's clothes, you don’t believe your expression of grief will bring her back, but you imagine her there
and wish she were. When you hide under your sheets in fear, you don’t believe it will keep you safe, but you wish it would, and you imagine it is. So, Velleman is able to explain such expressive actions by referencing imaginings and wishes even though such actions are a mix of symbolic and non-symbolic actions.

But here’s the important thing about Velleman’s account: it doesn’t save the standard account. We can explain such behaviors in terms of imaginings and wishes, but that is not, for Velleman, a “substitute” for beliefs and desires. Velleman’s account is supposed to be a rival of the standard account. Imagining and wishing for something does not rationalize the action in the way belief/desire pairs rationalize action. At most, “mock-beliefs” and “mock-desires” “mock-rationalize,” even for Velleman. And so, Velleman’s view is not going to help us here to save the standard account. I shall return to Goldie. In the next section, I’ll show that Goldie’s account of an imagining/desire pair cannot rationalize action either. And since rationalizing action is a necessary feature of the standard account, Goldie’s view does not in fact save the standard account from Hursthouse’s challenge.

**III. d. Imaginings Cannot Rationalize Action**

I’ve been working hard to limit the number of cases that might be explained by an imagining/desire pair. The number of cases is fewer than the number of symbolic expressive actions because many of the symbolic actions won’t be best explained by a “primitively intelligible” desire and an imagining. And obviously the other non-symbolic expressive actions won’t be explained by an imagining/desire pair either
since there is no imagination involved. So, we have at least three categories of expressive actions, (1) symbolic expressive actions explained by Goldie’s imagining/desire pair, (2) symbolic expressive actions not explained by an imagining/desire pair, and (3) non-symbolic expressive actions not explained by an imagining/desire pair. Only (1) answers Hursthouse’s challenge to the standard account. In this section, I will argue that even symbolic expressive actions explained by an imagining/desire pair do not save the standard account of intentional action from Hursthouse’s challenge. And so, Hursthouse’s challenge remains.

The standard account, according to Goldie, says that intentional actions “can, quite generally, be explained and understood as rational under a description by the attribution to the agent of a combination of some sort of pro-attitude, such as desire, and some sort of cognitive state, such as belief” (Goldie, 26). Such a rationalizing explanation gives the reason why the agent acted. Or in Hursthouse’s terminology, “According to the standard account of actions and their explanations, intentional actions are actions done because the agent has a certain desire/belief pair that explains the action by rationalizing it” (Hursthouse, 57). Goldie substitutes imagining for belief in the explanation of expressive actions, and he thinks this solves the problem. But why think imagining something—under any description—could rationalize action?

What rationalizes things? Well, in the most general sense, reasons rationalize things. What is it about reasons that rationalize? Presumably, reasons are considerations that count in favor (in some minimal sense) of doing or believing
something else. In the standard case, my desire for X and my belief that doing Y is the best way of getting X combine together to be the reason why I do Y. This belief/desire pair counts in favor of doing Y. Or sometimes I might have a perceptual experience of X, and this experience counts in favor of believing that Y. In this case, my perceptual experience of X is my reason for believing that Y. Perceptual experiences can be reasons for perceptual beliefs because such experiences are about the way things are (or appear to be). In the same way, beliefs are part of reasons for action because beliefs are about the way things are (or seem to be). So it seems that rationalization of an action must contain some element that represents the way things are.

There is a crucial difference between beliefs and perceptual states, on the one hand and imaginings on the other. Imaginings are fantastical, and by their very fantasy-orientated nature, they are in no way about how things are. Imagining something does not require the subject to consider that thing to have been the case. Believing or perceiving something does require the subject to consider that thing to be the case. So, imaginings differ from beliefs with respect to an attitude taken to the content of the mental representation. Beliefs require assenting to the way things are; imaginings do not.

It is because beliefs require the subject to make some connection between what they believe and the world that beliefs can rationalize further states and actions downstream. This is why beliefs and perceptions can be reasons—they represent something that is the case. When imagining something that is not the
case, there is no such connection required to what actually is the case. This is why beliefs can be reasons but imaginings cannot. Beliefs, perceptions, and other states that are about the way things are in the world can rationalize further states and actions. But imaginings—flights of fancy—simply aren’t suited to be the kinds of states that can rationalize anything. At most they can “mock-rationalize.”

So, we cannot just substitute any old cognitive state for belief and claim to have saved the standard account. Whatever cognitive state we combine with desire must itself be something that can be a reason, such as evaluative judgments, beliefs, perceptions, etc. Imaginings cannot be reasons. So, imaginings just aren’t the kinds of states that rationalize.

Even the few instances of (1) symbolic expressive actions explained by an imagining/desire pair do not rationalize action in the way required by the standard account. Imagining something to be the case simply does not rationalize action in any other kind of case, and so Goldie’s solution here is ad hoc. He presents imaginings to save the theory from a counterexample, but there’s no independent evidence for thinking imaginings can rationalize actions in other cases. In fact, there’s more reason to think imaginings do not rationalize anything—actions included. A fantasy cannot rationalize reality. So, Goldie has failed to save the standard account from Hursthouse’s challenge.
IV. Concluding Remarks

So far I’ve shown that attempts to save the standard account from Hursthouse’s challenge ultimately fail. What’s missing from Goldie and Velleman is reference to the emotion itself. What reason is there to refuse to consider emotions as rationalizing expressive action? Expressive action is by definition action expressing emotions, so why refuse to consider that the expressive action—expressing an emotion—is rationalized by the very emotion it expresses? Given the work from chapters 2 and 3, I believe emotions are the kinds of states that rationalize actions. They are about what is the case, and so they are different from imaginings and more like beliefs and perceptions.

In the next chapter, I consider two alternate solutions to Hursthouse’s challenge, from Sabine Doring and Monika Betzler. Both Doring and Betzler agree that the standard account cannot make sense of expressive actions. Both reject the standard account. And both think we must make some reference to emotion in explaining expressive actions. It seems as if we all agree! But, there are disagreements between my view and their views. I’ll present each of their views, and then I’ll raise some serious objections. Then I’ll sketch my own view, which I argue is the best explanation of expressive actions. I’ll provide a more thorough defense of my view in Chapter 6.
Works Cited


Chapter 5: Possible Alternatives to the Standard Account

I. Introduction

Jane, in a wave of hatred, scratches out Joan’s eyes in a photograph. Jane acts intentionally and deliberately. She storms towards the mantle, picks up the photograph of Joan, reaches for the scissors, and scratches out only Joan’s eyes. Jane is so angry because Joan has been having an affair with her husband. “You were supposed to be my best friend!” Jane shouts as she holds the torn photograph in her hands. What should we say about Jane’s action? The entire purpose of this chapter and the last is to make sense of Jane’s action (and others like it).

Here I’ll present two such attempts, one from Monika Betzler, and another from Sabine Doring. Both of these philosophers reject what I’ve called the standard account of intentional action:

**Standard Account:** For all actions, A’s action is intentional if and only if she does it for a reason where performing the action for a reason requires (1) A has some element in her subjective motivational set (usually desire) the satisfaction of which will be served by her φ-ing, (2) A has a belief about the means for satisfying her desire, and (3) A’s action involves a practical inference and not mere content-efficacy.40

---

40 More on the difference between practical inference and content-efficacy will follow in chapter 6.
The standard account is an account of how intentional actions are *rationalized* by other mental states of the agent.\(^{41}\) Betzler essentially replaces the belief/desire pair with an agent's ongoing evaluative perspective, while Doring argues that emotions can do all the work for beliefs and desires, making both means-end beliefs and desires unnecessary. I will argue that neither one of these alternatives to the standard account can satisfactorily explain Jane's action. Instead, I'll argue that the best explanation\(^{42}\) of Jane's action is that her anger with Joan combines with a desire to scratch her eyes out of the photograph. I think this is the simplest and most obvious proposal. The only supposed cost is revising the standard account. I'll argue this is no cost at all.

**II. Betzler's Account of Expressive Actions**

Betzler aims to show “that a core class of actions expressing emotions can be explained as rational, yet not within the framework of the standard model” (Betzler, 447). So far, so good. We agree. But Betzler attempts to explain expressive actions as rational by citing the agent's *ongoing evaluative perspective* as the reason for

---

\(^{41}\) I have framed the standard account as an account of rationalization, and not simply motivation. This will be important below. My reasons for doing so are that the standard account adheres to a Humean theory of motivation, which claims that desires are necessary for intentional action. But desires alone do not rationalize action—conditions (2) and (3) are also required for rationalizing the action.

\(^{42}\) I will use ‘explanation’ and ‘rationalization’ interchangeably. They both contrast with motivation, and I take both to be what the standard account is addressing. While ‘explanation’ might seem more causal than ‘rationalization,’ I do not think there is this difference. Both are causal.
which he acts. I'll first present what an ongoing evaluative perspective is, then explain how such a perspective gives an agent reasons to have emotions, and lastly how challenges to the evaluative perspective gives an agent reasons to act expressively. Then I'll raise a series of objections to Betzler's account.

Betzler's main thesis is that we cannot account for expressive actions by looking at the agent's momentary or isolated mental states such as beliefs and desires; instead we ought to consider a historical account. Such a historical account will revolve around the agent's ongoing evaluative perspective—a perspective he has acquired over time. The agent's ongoing evaluative perspective is formed as “a result of engagements that he has come to value over time and that shape the way he governs himself” (Betzler, 455). Such engagements might include important relationships with others, long-standing careers, projects, religious commitments, etc. These various engagements also become part of the agent’s own identity. As such, the agent feels a certain “pressure” to maintain his evaluative perspective so he can make sense of himself as an agent.

The agent's ongoing evaluative perspective gives him reasons to feel certain emotions. Betzler argues that “the agent perceives some object or state of affairs as calling for emotions like grief or jealousy only if he already enjoys a distinct relation to the object in question” (Betzler, 456). Jealousy represents a loss of affection from someone you care about. You wouldn’t feel as jealous of a stranger’s attention to someone else, when you don’t first care about the stranger’s attention. Without an evaluative perspective that includes caring for a person and her affection, you would
not feel jealous of the person showing affection elsewhere. If you didn’t first care about your career, then you wouldn’t feel angry when someone mocks it. If you didn’t first value a person’s life, then you wouldn’t grieve when the person passed away. The agent must’ve first valued the personal relationship—tying up his own identity with the relationship—as a prerequisite for feeling sad or jealous. And whatever becomes tied up with one’s identity in this way becomes part of the agent’s ongoing evaluative perspective. So, the agent’s ongoing evaluative perspective explains the agent’s emotions.

Not only does the agent’s evaluative perspective give him reasons to feel, it also gives him reasons to act. Betzler argues that “A loving relationship, then, yields not only reasons to act connected with the person one loves. It also yields reasons that govern one’s own life over time” (Betzler, 458). Here reasons get on the scene. Loving relationships—as part of one’s identity—give agents reasons to act in various ways. That is, one’s ongoing evaluative perspective issues in reasons to act. When you love your spouse, for instance, then that gives you reason to do certain things that make her happy, or things that are helpful to her, etc. Not only do the commitments tied up with one’s identity give an agent reasons to act, it also “generates reasons to sustain that perspective” (Betzler, 459). Having a particular evaluative perspective, then, gives an agent reasons to feel, reasons to act, and reasons to sustain that perspective over time.

But the real issue here is how expressive actions are to be explained as rational—without relying on the standard account. We’ve got to combine the way
evaluative perspectives give an agent a reason to act with how they give an agent reasons to feel—so we can see how it's rational for an agent to act expressively. Betzler reminds the reader that “an agent’s evaluative perspective, as he acquired it over time and thus identified with it, generates reasons to sustain that perspective” (Betzler, 459). This becomes especially important when there are challenges to the agent’s evaluative perspective—and thus to his identity. For Betzler, “the more it comes under pressure, the more he is liable to strong emotions that reflect his response to reasons for sustaining his perspective” (Betzler, 459). Then, the more he is liable to experience these strong emotions, the more powerful is the motivation to act expressively. Consider two expressive actions: kissing the dead beloved’s sweater out of grief, and the case of Jane scratching out Joan’s eyes in a photograph out of anger. As for the grief case, Betzler claims

“In caressing the clothes of a loved one who has passed away, we express our suffering from the loss of value that the death presents. We try to hold on to that person’s value for us by valuing objects that remind us of him” (Betzler, 460).

In the anger case, Betzler claims,

“Tearing apart a photograph of a person one is angry with is an expression of our disapproval of her behaviour. It is a civilized substitute for actually destroying that person, thus withdrawing our initial positive evaluation of her” (Betzler, 460).

These actions “motivated by such emotions are rituals that help to express or revise what one has come to value” (Betzler, 460).
Betzler offers several possible reasons why these ritual-like expressions of one’s emotions help to sustain one’s evaluative perspective. First, when the evaluative perspective is negatively challenged—when what we valued and identified with has gone away, as in the case of grief—we still need to demonstrate to ourselves that our evaluative perspective remains (Betzler, 460). Betzler claims, “we reinforce our evaluative perspective in acting expressively” (Betzler, 461). Second, sometimes the evaluative perspective is challenged because what we once valued no longer deserves our valuation, as in the case of Jane. Betzler argues we need to demonstrate to ourselves that the object deserves such devaluation, and this happens in expressive action. Third, an agent whose evaluative perspective has come under attack may be experiencing a “divided self.” Jane once valued Joan’s friendship, and now that friendship has come under attack. Betzler argues it is her “inner division that makes [her] prone to act expressively so as to gain coherence and stability” (Betzler, 461). For Betzler, the entire point of expressive actions is to sustain (or sometimes revise) one’s evaluative perspective by regaining control over it. She concludes:

“By having come to value engagements of the kind mentioned, however, he implicitly takes himself to have reasons to continue valuing what he values. And if this turns out to be impossible due to circumstantial constraints, he implicitly takes himself to have reasons to respond to that change” (Betzler, 462).

Here’s how I understand her account of expressive actions. Expressive actions express strong emotions, where these strong emotions are in response to one’s

---

43 I’m not sure how this answers the question. I’ll return to it in the objections below.
evaluative perspective being challenged by the environment. Since one’s evaluative perspective gives the agent reasons to have the strong emotions, then the evaluative perspective also gives one derivative reasons to express those emotions in ways that help to restore the agent’s ongoing evaluative perspective.

In the next section, I will raise a series of objections to Betzler’s account. In the end, I think that Betzler’s account must be understood as an instance of the standard account. So, it is no challenge to it. But even when we understand it as an instance of the standard account, it does not adequately make sense of Jane’s action. And so, I conclude that Betzler’s account does not explain or rationalize expressive actions at all.

III. Objections to Betzler

For Betzler, an agent acts expressively to sustain her evaluative perspective. The grieving man caresses the clothes of his deceased wife “to hold on to that person’s value for him by valuing objects that remind him of her.” That makes sense. He identified himself with this relationship for most of his adult life, and now she’s gone. He can’t suddenly stop thinking of himself that way or give up on valuing her. And in those moments when he is struck by an intense feeling of guilt, he sustains his evaluative perspective by caressing his deceased wife’s clothes. But what exactly is the connection between the evaluative perspective, the grief, and the caressing?

I’m tempted to interpret Betzler as saying that the man acts in order to sustain his evaluative perspective—in order to maintain some semblance of self. If
so, then he is no longer acting expressively at all. According to Hursthouse (the original articulation of expressive actions), expressive actions are defined as such:

“(i) that the action was intentional; (ii) that the agent did not do it for a reason in the sense that there is a true description of action of the form "X did it (in order) to..." or "X was trying to..." which will "reveal the favorable light in which the agent saw what he did," and hence involve, or imply, the ascription of a suitable belief; and (iii) that the agent would not have done the action if she had not been in the grip of whatever emotion it was, and the mere fact that she was in its grip explains the action as much as anything else does” (Hursthouse, 59).

I think Betzler’s account violates condition (ii)—on her view, it seems “expressive” actions are done in order to sustain an agent’s evaluative perspective. Betzler repeats many times that we have reasons to sustain our evaluative perspectives. The grieving man has a reason to continue valuing his late wife, even though she’s no longer in a relationship with him. So, he has reasons to perform certain actions that continue to preserve his “relationship” with her. One of these actions is caressing her clothes. Another way to put this is that he has a reason to caress her clothes in order to maintain his evaluative perspective. I don’t think Betzler has ruled out this interpretation of her view.

Betzler claims an anonymous referee raises a similar worry. The proposal could be recast within the belief/desire model in the following way. The agent has a desire to maintain a coherent ongoing evaluative perspective, and the belief that this is valuable, or the belief that acting expressively will help to maintain that evaluative perspective. I’m not sure of Betzler’s response to this. She states, “the adaptive function of coherence and stability is only served if the agent in question acts
according to other reasons than his desire for coherence” (Betzler, 464).\footnote{She doesn’t address the correctness or incorrectness of the beliefs.} She does not say what these “other reasons” are. My best guess is this is supposed to parallel the paradox of hedonism. If you pursue happiness itself, you won’t be happy. Instead, you have to pursue other enjoyable things—flying a kite or reading a book—and then the byproduct will be happiness. Maybe the same goes for evaluative perspectives. You can’t pursue coherence itself, or you’ll end up dreadfully incoherent. Instead, you have to pursue the things you value—relationships, projects, career, etc—and then you’ll just be coherent.\footnote{Again, this is all speculation. Betzler unfortunately spends only a paragraph responding to what seems to be the central objection to her view.} So, while an agent has reasons to sustain his evaluative perspective (something Betzler repeats time and time again), the agent will not respond to those reasons by having a desire to sustain it.

First, I’m not sure agents lack a desire to sustain their evaluative perspectives. I greatly value my relationship with my daughter. Our relationship has become part of my evaluative perspective—I identify myself through this relationship. When she and I have a disagreement (which happens quite often with teenagers), I want to resolve the argument because I also want to preserve our relationship. Most people desire to be their true—coherent, unique—selves. And we work very hard to achieve that goal. We rationalize away inconsistencies; we deceive ourselves into coherence. We go to long lengths to sustain our evaluative perspectives because we value a coherent evaluative perspective.
Second, I’m not sure an agent needs a desire for a coherent and stable evaluative perspective to get a belief/desire pair off the ground. A different pair would be that the agent desires to caress his deceased wife’s clothes, and he believes that doing that will help sustain his evaluative perspective—it will help him maintain his own identity. I think this pair is much closer to Betzler’s view. What else might she mean by saying it is the agent’s “inner division that makes him prone to act expressively so as to gain coherence and stability” (Betzler, 461). The agent believes that caressing the clothes will help him gain coherence and stability, and this combines with a desire to caress the clothes to explain his action. At this point, all Betzler could mean by “acting expressively” is “those examples Hursthouse gave.” If the agent is acting in order to (“so as to”) gain coherence and stability, then he is no longer acting expressively.

Third, Betzler’s account would also violate Hursthouse’s condition (iii): that the agent would not have done the action if she had not been in the grip of whatever emotion it was, and the mere fact that she was in its grip explains the action as much as anything else does. Betzler’s account seems to be exhausted by the belief/desire pair. The emotion is simply a byproduct of the challenge to the evaluative perspective. As such, it’s not doing any real explanatory work. The evaluative perspective (and challenges to it) gives the agent reasons to feel and reasons to act. It gives an agent reasons to act in the belief/desire way—the agent desires to do something and has a belief that doing it will preserve his evaluative perspective. It also gives an agent reasons to feel—the agent values his relationship with his wife,
and this gives him reason to feel sadness at the loss of that relationship. But this is a point with two vectors—the action vector and the feeling vector. But what Betzler would need to satisfy Hursthouse’s condition (iii) would look like this:

   Evaluative perspective $\rightarrow$ emotion $\rightarrow$ expressive action

This does not seem to be her view.

Lastly, I do not think the same reasons to sustain one’s evaluative perspective could be used to explain Jane’s action. The grieving man is acting to sustain his evaluative perspective; Jane is acting to destroy her identity as a friend of Joan’s. According to Betzler, “Tearing apart a photograph of a person one is angry with is an expression of our disapproval of her behaviour. It is a civilized substitute for actually destroying that person, thus withdrawing our initial positive evaluation of her” (Betzler, 460). The first thing to note is that this explanation makes no reference to Jane’s evaluative perspective. I disapprove of Joan’s behavior. One ought not have an affair with one’s best friend’s husband. It’s quite despicable. I could also act in ways that express my disapproval. I might not go so far as scratching her eyes out in a photograph, but I might shun her from possible future activities, even though I don’t currently have any relationship with her. And in disapproving of her behavior and shunning her, I am thereby withdrawing any positive evaluation of her. But none of this requires referencing my evaluative perspective. I never had a close relationship with Joan. That is, I could satisfy Betzler’s explanation of Jane’s action, and yet a relationship with Joan is not part of my evaluative perspective.
The second thing to note about Jane is that even if there is some implicit reference to her evaluative perspective, she is surely not acting in order to sustain it. She is scratching the eyes in the photograph (most likely) in order to symbolically destroy any value she once placed on Joan or her relationship with her. Even according to the belief/desire explanation, Jane might have a desire to scratch the eyes out in the photograph, but she does not do this with the belief that so doing it will sustain her evaluative perspective. It’s quite the opposite.

Betzler’s account of expressive actions, then, is unsatisfactory. It’s not an account of expressive actions at all insofar as the agent acts in order to sustain his evaluative perspective (consistent with the standard account), and insofar as emotions play no explanatory role. But even if Betzler were to backtrack and say it was supposed to be an account of actions-like-Jane’s, sustaining one’s evaluative perspective plays no role in explaining Jane’s action. I think we should look for a different solution.

IV. Doring’s Account of Expressive Actions

Sabine Doring also argues, in “Explaining Action by Emotion,” that the standard account is inadequate. Doring’s main conclusion is that the emotions constitute an irreducible category in the explanation of action (Doring, 214). Accordingly, emotions by themselves can rationalize action—in the absence of desires and means-end beliefs. She thinks there are two ways emotions can explain action: (1) there is expressive explanation of action, and (2) there are actions that are not only
explained, but also rationalized by the emotion’s intentional content (Doring, 214). When actions performed out of an emotion are rational, the emotion’s rationalizing capacity is attributed to its representational content (Doring, 215). If the emotion’s representational content provides an end for action, then the action is rationalized. In expressive cases, emotions provide no such end—only motivation. This would imply that all expressive actions are arational. But expressive actions and rational actions are both motivated actions. The motivational aspect of emotions comes from the emotion’s affect—the way it feels to be sad, angry, afraid, jealous, etc. Even though the representational form of emotions may change, all emotions motivate. With this contrast in hand, I’ll present Doring’s view in more detail.

For Doring, all emotions motivate in the absence of desires. Doring thinks the problem with Goldie, Smith, Raz and the like is that they each presume the so-called Humean theory of motivation, according to which all motivation is ultimately due to the desire an agent actually has (Doring, 220). But for Doring, emotional motivation is independent of desire, that is, independent of whether the emotion provides an end for action (Doring, 215). Instead, the motivational force of an emotion has to be explained by reference to the emotion’s affect. She argues, “an emotion’s motivational force cannot be reduced to having a world-to-mind direction of fit, and therefore not to desire, but has to be understood in terms of what I shall

---

46 I disagree. I think expressive actions can be rationalized by emotions as well.
call ‘affective perception’” (Doring, 220). Such reference to an emotion’s affective component is required in the explanation of action by emotion (Doring, 223). Unlike a desire, an emotion’s affect can still move its subject to act even if it is not necessary or actually impossible to change the world in such a way that it fits the emotion. Someone might grieve a loved one without doing anything to bring her back; and for Jane, “the representational content of Jane’s emotion is that Joan is an awful person. Seeing Joan as an awful person does not by itself provide an end for action” (Doring, 224). This would be the expressive case—Jane’s anger towards Joan does not set an end for action when she scratches the eyes out of the photograph. And yet, the way Jane feels when she’s angry nonetheless motivates her action, just not in a desire-like way.

But Doring contrasts such “aimless” expressions of emotions with the more rational type. In the expressive case, Jane’s anger represents Joan as an awful person, where she can’t do anything to change that. But consider a different case. Imagine Jane is walking towards Joan in a very aggressive way. According to Doring, “Joan’s fear of the aggressive-looking woman has a representational content that provides an end for action, namely, the end of avoiding danger” (Doring, 224). When Joan crosses the street to avoid Jane, her action is rationalized by her fear, whereas when Jane scratches the eyes out of the photograph, her action is merely motivated by anger but not rationalized by it. Doring then concludes that whether

---

47 Doring does not give an account of how an emotion’s affect—in the absence of providing an end for action—could motivate action at all. This is a problem for her view.
an action springing from an emotion can adequately be explained in terms of means-end reasoning depends on that emotion’s representational content (Doring, 224). Having a representational content, an emotion essentially has a mind-to-world direction of fit. Only in those cases where an emotion at the same time has a world-to-mind direction of fit does it supply a reason for acting, and it does so only in a mediated way. Yet, even when emotions have world-to-mind directions of fit, that is not a state called “desire.” So, emotions can motivate in the absence of desire because of their affect, and when such affect is combined with a particular kind of representational content—the kind with a world-to-mind direction of fit—the emotion also rationalizes action.

In the following section, I’ll raise two objections to Doring. First, Doring does not give an account of how the affective component of emotions motivates action, and I’ll argue that the way an emotion feels can at most prepare one for action, but it cannot alone motivate action. Second, Doring argues that the difference between expressive and rational actions lies in a difference in the representational content of an emotion. But if the representational content of all emotions is an evaluation of its target, then the representational content alone cannot differentiate between expressive and other actions. Instead, we’d need some other state—a goal-directed one such as desire—to combine with this representational content to fully rationalize action.
V. Objections to Doring’s Account

V. a. Affective Motivation—Is There Such a Thing?

I’m not sure how Doring’s account here undermines the Humean theory of motivation (the desire part of the standard account). According to most theories of motivation, what’s essential to motivation is a state that is *action guiding*. Such a state must be directed at bringing about goals of the subject. Only desires satisfy these requirements for the Humean. Doring thinks emotions can also satisfy these requirements of motivation. But can the affective, or feeling, component of emotions really satisfy these requirements? Imagine fear. Suppose there is a burglar trying to get into my home. I see him try to jimmy the lock, fail, and then use a brick to break the window. He is reaching around to unlock the door. How am I feeling right then? I’m sure the hairs are standing straight up on the back of my neck. I’m probably sweating, pale, and my heart’s surely racing. I’m standing there in the kitchen with a knife, poised for a fight, desperately hoping the police show up soon. Do these feelings alone—in the absence of any desires—motivate me to lunge towards the intruder as he comes through the door? Suppose we can take that moment—just seconds before the intruder makes his way into the kitchen—and isolate my mental states. Let’s abstract away all of my desires—desires for safety, desires for protecting my home and family, even the most basic desire to hold the knife in my hand. All that’s left is my representation of the man as dangerous and the pulsing sensation of fear running through my veins. Can the way the emotion feels in my body guide me through a series of complex movements? I don’t think it
can. The affective component of the emotion isn’t even intentional—it’s a mere sensation. It’s not about anything. It cannot move my hand in this way and that to threaten the robber.

Surely the strong feelings in my body play some role in action. Everyone would agree to that. But the point I’m challenging is that they play a motivational, action-guiding role. I think that would be attributing too much to such sensations. Instead, I think the bodily feelings prepare me for action—they get my blood flowing. But they do not then determine how I move my body in a complex action. They take an action such a lunging and intensify it. The fear sometimes gives people “super strength.” They might also delimit my available actions. For example, as my body is pumped full of adrenalin, such feelings close off the possibility of heading to the living room to watch some TV. But do the feelings guide my behavior? Do they set a goal for me? I don’t think so.

For Doring there are two requirements that must be met for motivation by emotion: you must classify the intruder as fearsome, and your physiological economy must be properly mobilized for flight or fight. The classification of an intruder as fearsome is not obviously felt at all—it is representing the world to be a certain way. And the body being prepared to fight or flee is necessary for action, for sure, but it does not encompass all of motivation. Body preparedness is just that—a feeling in the body. The body is primed to act. But this alone cannot direct the agent’s action. Body preparedness cannot point the agent towards the intruder, stick out a knife, and threaten him. It merely prepares the body to perform some
further action or other. And so, it seems Doring’s account of motivation is no rival to the Humean one because it’s not an account of motivation at all. It is an account of preparedness to act; it is not an account of motivation to act.

But the problem here runs even deeper. If Doring maintains that all emotions are motivational, and yet only some actions done out of emotions are also rationalized, then it seems her view won’t quite challenge the standard account. For any account, motivation is necessary for action. By Doring’s own lights, motivation comes in two forms: Humean and non-Humean. Humean motivation is desire, or states with world-to-mind directions of fit. Doring’s non-Humean motivation comes from an emotion’s affective component, which need not have a world-to-mind direction of fit. The standard account of action accepts only Humean motivation as a reason for action; only states with world-to-mind directions of fit can rationalize action. Doring is attempting to argue against the standard account. To do so, she’d have to show how non-Humean motivation still counts as a reason for action. But her distinction between acting for a reason and not acting for a reason relies on Humean motivation, such that acting for a reason requires states with a world-to-mind direction of fit (desires).

There are at least three options here: (1) if expressive actions are done for a reason, then the reason must involve Humean motivation, (2) expressive actions can be done for a reason, but the reason involves non-Humean motivation, (3) if expressive actions involve non-Humean motivation, then they are not done for a reason. Doring accepts (3), but the only option that is contrary to the standard
account is (2). The standard account accepts both (1) and (3)—being done for a reason “rises and falls” with Humean motivation. So, Doring’s account of reasons for action is in line with a standard account.

But maybe we should backtrack a bit here. We could re-cast Doring’s view: she’s trying to give an explanation of expressive actions, for sure. But where her real work is done is by contrasting such arational expressive actions with other actions that are rationalized by reference to emotions, and emotions alone. And it is this latter class that challenges the standard account more directly. However, I don’t think this solution will work. I don’t think it will work because I don’t think we can sustain the distinction between expressive and “rational” actions done out of an emotion. I’ll turn to that objection next.

V. b. Doring’s Distinction Between Expressive and Rational Actions Won’t Work

Consider again the difference between Jane and Joan. Jane is angry with Joan for having an affair with her husband. She is enraged. From all this anger, she scratches out Joan’s eyes in a photograph. For Doring, Jane’s action is motivated by her anger, but her anger does not rationalize her action. The contrast case is from Joan’s perspective. Suppose Jane is approaching Joan in a very aggressive way. Joan knows what she’s done, and she now suspects Jane knows, too. As she sees Jane getting closer and closer, she becomes very afraid of what Jane will do to her. From all this fear, she starts running in the other direction of Jane. For Doring, Joan’s action is motivated by her fear, but it is also rationalized by her fear. What’s the
difference? For Doring, Joan’s fear sets an end for action whereas Jane’s anger does not.

But why? I think Doring’s account here is putting the cart before the horse. The cart is supposed to be the emotion—driving the action (the horse). We are trying to use the emotions to explain, motivate, and rationalize action. Emotions are supposed to be playing the role of belief/desire pairs in action. But Doring is taking some aspect of the action itself—whether the action actually affects the target of the emotion—to alter an assessment of whether the emotion is setting the end for action. The difference between Jane and Joan is that Jane is doing something that doesn’t change anything with Joan, whereas Joan is doing something that does affect Jane’s chances of catching up with her and harming her. We can’t simply read off of what their actions were to whether the emotions—both reasonable emotions—set an end for action.

If we compared two instances of anger, instead of comparing anger with fear, my point should be clearer. Suppose we compare Jane, who out of anger scratches out the eyes in the photograph with a different account of Joan. After scratching the eyes out of the photograph, Jane marches over to Joan’s house with a crowbar and completely demolishes Joan’s car. She causes thousands worth of damage. Upon discovering this, Joan becomes very, very angry with Jane. As a result, Joan calls the police. So, Jane is angry with Joan for having an affair with her husband, and she scratches her eyes out in a photograph. Joan is angry with Jane for bashing up her car, and so she calls the police. Doring is asking us to consider the fact that calling
the police will have an actual effect on Jane, whereas scratching up the photograph will have no such effect on Joan as evidence for thinking there is something fundamentally different about Joan’s and Jane’s anger. Jane’s anger does not set an end for action whereas Joan’s does.

This is a mistake. Jane’s anger has the same form as Joan’s. Jane has wronged Joan, and Joan has wronged Jane. Both of their anger is justified. Both of their anger represents the target as having wronged them. Both of their anger feels a certain way—their hearts are racing, their jaws are clenched, they’re both making fists with their hands, their muscles are contracted throughout their bodies. Their anger prepares both of them for action. The only difference is which action they take. Jane scratches up a photograph and Joan calls the police. What is the best explanation of this difference in action? I don’t think the answer lies in some fundamental difference in their anger—that’s the same for both of them. Instead, I think the fundamental difference lies in their desires. Jane wants to scratch up a photograph, and Joan wants to call the police. It’s as simple as that. We might say that Joan’s desire is better or more reasonable or more effective at achieving some other goal she has, but nonetheless, the most obvious difference between Jane and Joan lies in their different desires and not their similar emotions. Not only is this the most obvious difference, I think it is the correct explanation of the difference between Jane and Joan. In the following section I will present my own account of actions expressing emotions. I will accept the idea that desires are necessary for rationalizing action—expressive or otherwise—but I will challenge the standard
account by also requiring emotions as states that are sometimes necessary for rationalizing action as well.

VI. Emotions Must Play a Role

I believe that desires set the goal for action; they set the end for which an agent acts. The agent does act intentionally in an expressive action, and some psychological state must explain the action. The psychological states that explain the action are the agent’s emotion combined with a desire to set the end for action. My view in its most basic form is that this emotion then combines with a desire to rationalize the action. There may be many other beliefs and desires “on the scene” when Jane acts. She surely believes that the photograph on the mantle is a picture of Joan. She believes Joan had an affair with her husband. She desires to scratch out the eyes in the photograph. She may also want her husband to walk in and see the mutilated photograph. Beliefs and desires are all over the place. The problem is—any candidate belief/desire pair is insufficient to rationalize Jane’s action. Her anger must be present for any of the candidate actions to make sense. I think this is a strong reason for thinking her anger combines with different desires to rationalize different actions. And in such cases, the emotion is necessary to rationalize the action. This is how my view differs from the standard view. The standard view (typically understood) claims that only beliefs and desires are necessary for rationalizing action, and that beliefs and desires are sufficient for rationalizing all

---

48 I argued for this conclusion extensively in the previous chapter.
intentional actions. I argue in the next section that sometimes beliefs and desires are insufficient for rationalizing certain actions, and for those actions, reference to an emotions is necessary.

**VII. Emotions Play a Rationalizing Role**

I take the concept “rationalization” to be a genus with several subspecies. A qualifying species of rationalization must be able to show how an agent’s action makes sense. It must meet the intelligibility requirement, as it’s often called. The familiar subspecies of rationalization comes from Davidson. This account says that each intentional action is rationalized by joining a desire for some end with a belief about the means to achieving that end. That is, the familiar schema for rationalizing action is through means-end reasoning. The schema is a good one because it meets the intelligibility criterion for all rationalizations. Citing that the agent desires X, and then combining this with a belief about how to get X is said to make perfectly good sense of an agent’s action.

A second subspecies has roughly the following schema: take desires combined with means-end beliefs, and then add some further condition (typically something trying to capture rational agency) on top. Take Bernard Williams, for example. He thinks more than a desire combining with a means-end belief is required. According to Bernard Williams in “Internal and External Reasons,” someone has a reason to \( \varphi \) in circumstances C if and only if she would desire that she \( \varphi \)'s in circumstances C if she were fully rational. To rationalize an agent’s action,
then, requires an idealization of the agent’s mental states. It’s not the actual states of the agent that rationalize action; it’s what the agent would desire if she were fully rational. Michael Smith offers a similar account of rationalization in *The Moral Problem*. He argues that “to say that we have a normative reason to \( \varphi \) in certain circumstances \( C \) is to say that, if we were fully rational, we would want that we \( \varphi \) in \( C \)” (Smith, 181.) This analysis goes far beyond requiring only actual desires and means-end beliefs. It requires the beliefs and desires of an idealized, fully rational agent. David Velleman also adds additional requirements for rationalization. He thinks we have desires and means-end beliefs, and these are sometimes reasons for action. But to *rationalize* action, the agent must also have “a desire to act in accordance with reasons” (Velleman, *The Possibility of Practical Reason*, 141). While there’s disagreement about how exactly this subspecies of rationalization shakes out, there is good reason to suppose that some proper rationalizations do require something in addition to actual means-end beliefs and desires of an agent. In the one case, it’s the beliefs and desires of an ideally rational agent that count, in the other it’s the beliefs and desires of an agent who has the second-order desire to act in accordance with reasons. Either way, some extra condition is added to Davidson’s original account—the subspecies I first presented. So, some subspecies of “rationalization” require more than what the simpler standard account requires.

The third subspecies of rationalization I’d like to consider comes originally from Anscombe’s *Intention*, and it is an *expressive* sense of rationalization. Anscombe claims, “very often, when a man says ‘I am going to do such-and-such’, we
should say that this was an expression of intention” (Anscombe, 1). Sometimes when I say “I am going to take a walk,” this expresses my intention to take a walk. Such an action is rationalized by expressing my intention. My action is rationalized by my intention, which is expressed in action.⁴⁹ Elizabeth Anderson presents a similar expressive account in *Value in Ethics and Economics*: “an expressive theory defines rational action as action that adequately expresses our rational attitudes toward people and other intrinsically valuable things” (Anderson, 17). Such an expressive theory meets the intelligibility requirement. When performing some action, say, taking your child to school, your action makes sense by referencing simply your intention to take your child to school. Or, it makes sense by referencing your rational attitude of valuing your child’s education.⁵⁰ By expressing intentions or other rational attitudes, we make sense of action. When action makes sense in this way, then we can say the action is rationalized.⁵¹

The goal of this paper is to defend a further subspecies of rationalization. I claim that actions can be rationalized by a desire, which sets the end for action, combined with an emotion, which represents a relational property between the agent and the world. Actions then express this emotion/desire pair, similarly to how actions expressed intentions or values above. The difference is that actions can

---

⁴⁹ I should note that expressing an intention in this way is non-instrumental. The agent isn’t acting in order to expressing his intention; the action just is an expression of his intention.

⁵⁰ Again, this is a non-instrumental sense of rationalization. You are not taking your child to school in order to satisfy your values or goals of education; taking your child to school is just an expression of your values.

⁵¹ The conclusion here may be contentious, but I hope it’s intuitive that there are several different ways we make sense of actions.
be rationalized by expressing emotions and desires, whereas none of the other subspecies above could say that.

Emotions represent core relational themes—those things that matter to us as human beings. Here’s a list of the core relational themes represented by each major emotion, as first stated by Lazarus:

<table>
<thead>
<tr>
<th>Emotion</th>
<th>Representation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anger</td>
<td>represents a demeaning offense against me and mine</td>
</tr>
<tr>
<td>Anxiety</td>
<td>represents facing uncertain, existential threat</td>
</tr>
<tr>
<td>Fear</td>
<td>represents facing immediate, concrete, and overwhelming physical danger</td>
</tr>
<tr>
<td>Guilt</td>
<td>represents having transgressed a moral imperative</td>
</tr>
<tr>
<td>Sadness</td>
<td>represents having lost something valuable</td>
</tr>
<tr>
<td>Shame</td>
<td>represents having failed to live up to an ego-ideal</td>
</tr>
<tr>
<td>Envy</td>
<td>represents wanting what someone else has</td>
</tr>
<tr>
<td>Jealousy</td>
<td>represents resenting a third party for loss or threat to another’s affection</td>
</tr>
<tr>
<td>Disgust</td>
<td>represents taking in or being too close to an indigestible object or idea (metaphorically speaking)</td>
</tr>
<tr>
<td>Happiness</td>
<td>represents making reasonable progress toward the realization of a goal</td>
</tr>
<tr>
<td>Pride</td>
<td>represents enhancement of one’s ego-identity by taking credit for a valued object or achievement, either one’s own or that of some group with whom we identify</td>
</tr>
<tr>
<td>Relief</td>
<td>represents a distressed goal-incongruent condition that has changed for the better or gone away</td>
</tr>
<tr>
<td>Hope</td>
<td>represents fearing the worst but yearning for the better</td>
</tr>
<tr>
<td>Love</td>
<td>represents desiring or participating in affection, usually but not necessarily reciprocated</td>
</tr>
<tr>
<td>Compassion</td>
<td>represents being moved by another’s suffering and wanting to help</td>
</tr>
</tbody>
</table>

Table 1: Core Relational Themes
These are relational properties. A loss is only a loss to me of something in the world that I valued. I am one side of the relata, and the thing valued is the other side. All such relations represented by emotions are between the world and myself. And all such relations involve some assessment of how the subject is faring in the world. A loss signals the subject is faring poorly; an enhancement of one’s ego-ideal signals the subject is faring well.52

Jane—in a wave of jealousy—scratches out Joan’s eyes in a photograph because of her affair with Jane’s husband. In this case, Jane’s emotion combines with a desire to scratch the eyes out of the photograph to rationalize Jane’s action.53 Jane resents Joan for taking away her husband’s affections. And Jane’s jealousy here is absolutely appropriate—Joan has taken away his affection. Because of such jealousy (combined, I’m sure with rage), Jane scratches out Joan’s eyes in the photograph. The jealousy, I argue, is necessary to rationalize Jane’s action. Simply stating a desire to scratch the eyes out combined with a means-end belief that moving her hands in this-and-that way is inadequate as a rationalization. It doesn’t give us any sense of why Jane is doing what she’s doing. Only when we explain that Jane is enraged by Joan’s actions can we rationalize—and make sense of—what she’s doing. Considering actions such as Jane’s, then, ought to convince us there’s a

52 This account of emotions was developed more fully in chapter 2.
53 Jane might act in a different way that some might think makes more sense—she might instead try to have an affair with Joan’s husband, just to get back at her. Nonetheless, I think there is something to Jane’s scratching out Joan’s eyes in the photograph that does make sense to us—she’s very angry with Joan, but she deep down doesn’t want to go so far as to physically harm the actual Joan, so she symbolically harms Joan instead (but still as an expression of her anger and not in order to express her anger). It might turn out to be the more reasonable thing to do, given other of Jane’s values, beliefs, and desires on the whole.
different kind of rationalization here. But to have a full account of this emotion rationalization, it seems there should be a broader characterization. It is not enough to go case-by-case pointing out an emotion as a rationalizer. We need some sort of “schema”—something that will make sense over a wide variety of cases.

What would such a schema look like? Well, any such schema will inevitably be vague here. We have clear cases at the extremes. At one extreme, Jane’s action seems to be best rationalized by her emotion. But what if, out of anger, I smack my lips twice and click my tongue. You ask me why I did this, and I say “because I was so angry with Tom!” Does my anger here make any sense of my action? It seems not. So, at the other extreme we have nonsensical expressions of emotions. So, not every action that is caused by an emotional state will be rationalized by the emotion. But we don’t yet have anything general enough to count as a rationalization schema.

I suggest we look at the typical “action tendencies” of each emotion. Anger, for instance, tends toward retaliation. When Jane scratches out the eyes of her offender in a photograph, this action looks enough like retaliation—even if symbolic retaliation—to make her action intelligible to us. When I smack my lips twice and click my tongue out of anger, this has absolutely nothing to do with retaliation—symbolic or otherwise. There’s not even an object that is being hurt or kicked or torn or destroyed in some way that looks like what happens in anger. There’s not

---

54 It might be more intelligible if she went out and tried to scratch out Joan’s actual eyes—actual retaliation. But scratching the eyes out of the photograph is symbolic of actual retaliation, and so it makes some sense. And the only way it makes any sense at all—which I assume it does—is by referencing her anger.
even a person or thing that my action is aimed at. It looks quite strange to do this out of anger. I think we can give similar accounts of each of the core relational themes stated above. While there will inevitably be vagueness here, I hope we can go far enough to have a real idea of which actions out of emotions will make sense and which ones won’t. Here are my proposed action tendencies for each of the emotions listed above:

<table>
<thead>
<tr>
<th>Emotion</th>
<th>Tendency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anger</td>
<td>Tends toward retaliation</td>
</tr>
<tr>
<td>Anxiety</td>
<td>Tends toward withdrawing from the uncertain, existential threat</td>
</tr>
<tr>
<td>Fear</td>
<td>Tends toward a fight or flight response</td>
</tr>
<tr>
<td>Guilt</td>
<td>Tends toward reconciliation or sometimes concealment</td>
</tr>
<tr>
<td>Sadness</td>
<td>Tends toward slower movements, crying, intense remembering and reverence for the thing lost</td>
</tr>
<tr>
<td>Shame</td>
<td>Tends toward concealment</td>
</tr>
<tr>
<td>Envy</td>
<td>Tends toward trying to get the thing wanted that someone else has.</td>
</tr>
<tr>
<td>Jealousy</td>
<td>Tends toward trying to hold on to the loved one's affection and retaliation towards the third party</td>
</tr>
<tr>
<td>Disgust</td>
<td>Tends toward distancing oneself from the indigestible object or repulsive idea</td>
</tr>
<tr>
<td>Happiness</td>
<td>Tends toward behaviors that display hope for the future</td>
</tr>
<tr>
<td>Pride</td>
<td>Tends toward boasting</td>
</tr>
<tr>
<td>Relief</td>
<td>Tends toward pushing forward with a project</td>
</tr>
<tr>
<td>Hope</td>
<td>Tends toward trying to make the future better</td>
</tr>
<tr>
<td>Love</td>
<td>Tends toward sharing in the joys and pains of the loved one, as well as helping to improve the loved one's life</td>
</tr>
<tr>
<td>Compassion</td>
<td>Tends toward helping someone else</td>
</tr>
</tbody>
</table>

Table 2: Action Tendencies

I use “tends toward” here because some emotions might be expressed symbolically—as Jane’s was—and other emotions might not be expressed at all.
because the agent has suppressed any outward expression.\textsuperscript{55} The general schema would be that a desire combines with an emotion to explain action. And when the action expresses the emotion in a way that resembles the relevant action tendency above, then the emotion also rationalizes the action. When I smack my lips twice and click my tongue, such an action is caused by my anger, and I even cite my anger as the reason I did it, but such an action is not rationalized by my anger. It’s not rationalized by my anger because there’s nothing whatsoever that looks like retaliation in my action. Such an expression of anger simply doesn’t make sense.

Consider another example. Suppose I am disgusted by the piece of fudge shaped to look like feces. Even though I know it is just fudge, I still refuse to eat it. I might even push my chair away from the table or turn my head away from the lump. These would all be rational expressions of my disgust, and because of my disgust, such actions make sense. They make sense—given my disgust—even though I also believe that it’s only fudge and not actually feces. But what if instead I express my disgust by getting up to hug the experimenter? That wouldn’t make any sense at all! She just walked in here to serve me what looks like feces—feces I’m disgusted by. Disgust prompts us to retreat from the contaminated object, as well as whoever brought such a contaminated object to us. A hug would be completely antithetical to

\textsuperscript{55} I should also note that these “action tendencies” are based primarily on empirical data. There may be something about anger itself that tends toward retaliation, but when it comes to appropriate expressions of such retaliation, that depends on how people in a certain society often retaliate. That is, this rationalization of action will inevitably be relative to one’s culture.
what disgust tends to get us to do. And so, such a hug would make no sense whatsoever.

Given all this, I think we can say more generally how emotions rationalize action.

**Emotion-Based Rationalizations:** For some actions, A’s action is done for a reason if and only if (1) A has some element in her subjective motivational set (usually desire) the satisfaction of which will be served by her φ-ing, (2) A has an emotion that represents a relation between herself and something she values in the world, and (3) A’s action falls within the normal range of action tendencies for the relevant emotion.

I grant that the “normal range of action tendencies” will inevitably be vague. We don’t all express our emotions in the same way, for sure. But that doesn’t mean there are no clear cases. When someone offends me deeply, and I yell at him out of anger, my reaction is rationalized by my emotion. When someone offends me deeply, and I spin around on my head out of anger, my reaction makes no sense whatsoever because it is too far outside the normal range of action tendencies for anger. When I achieve some important goal I’ve set for myself, and I go for drinks with friends to celebrate, my decision seems to be rationalized by my pride in my achievement. But if instead, I hid my face, alone, in a closet “because I was so proud of myself!” my bizarre action would not be rationalized by pride. It looks more like shame. I think most of us are very skilled at deciphering actions that are rationalized by emotions from those that aren’t.
VIII. Concluding Remarks

In the previous chapter, I presented several attempts to save the standard account from the challenge of expressive actions. And I found each attempt to be inadequate, leaving me to conclude the standard account could not be saved. In this chapter, I presented three different accounts of expressive actions that are proposed as rivals to the standard account. I found both Betzler’s and Doring’s accounts to be unsatisfactory. Betzler didn’t actually give an account of expressive actions—only actions done out of an evaluative perspective. Doring’s account only gave emotions a motivating role, and not a rationalizing role. Since the standard account is an account of rationalizations, Doring’s account ended up being in line with it.

Finding all of these accounts of expressive actions deficient, I turned to the most obvious explanation: the emotion rationalizes the action, combined with desire. In the following chapter, I will expand my account of expressive actions here to cover more non-expressive actions. I will argue that not only do emotions rationalize actions, emotions can be the reasons for which one acts. That is, an agent can act for a reason by acting for an emotion.
Works Cited


Chapter 6: Emotions as Reasons

I. Introduction

In the previous chapters, I considered the case of Jane, who in a wave of anger scratched her nemesis’ eyes out of a photograph. Expressive actions such as this were put forth by Hursthouse as a challenge to the standard account of intentional action:

**Standard Account:** For all actions, A’s action is intentional if and only if she does it for a reason where performing the action for a reason requires (1) A has some element in her subjective motivational set (usually desire) the satisfaction of which will be served by her φ-ing, (2) A has a belief about the means for satisfying her desire, and (3) A’s action involves a practical inference and not mere content-efficacy.\(^{56}\)

Supposedly, Jane’s action does not meet condition (2), and yet her action seems intelligible to us. Hursthouse claims it’s understandable only by reference to her anger. In chapter 4, I considered several possible explanations of Jane’s action—each meant to save the standard account—and I found them all inadequate. Raz, Smith, Goldie, and Velleman all required a desire to rationalize action. They all accepted a Humean theory of motivation, condition (1) in my formulation. They all

\(^{56}\text{More on the difference between practical inference and content-efficacy will follow.}\)
agreed that the means-end belief in condition (2) was not met. However, they each tried to add different states to the Humean theory of motivation to save the standard account of intentional action. Smith tried to add emotion to the belief/desire pair, and Goldie and Velleman tried to substitute imaginings for beliefs altogether. I agree desires are necessary for motivation. So, I accept condition (1). But I think Smith is wrong to simply tack on emotions to belief/desire pairs. And I think Goldie and Velleman are mistaken in replacing beliefs with imaginings. Where might that leave me? I believe that emotions can combine with desires to rationalize action—expressive or otherwise. This view accepts the Humean theory of motivation, but it revises the standard account of intentional action. We should sometimes “substitute” an emotion for belief in order to properly explain an intentional action. I believe that the best explanation of Jane’s action is that her anger combines with a desire to scratch the eyes out. That is, emotion/desire pairs can sometimes rationalize expressive actions where the belief/desire “standard account” fails. This led me to what I called the “Emotion-Based Account” of expressive actions:

**Emotion-Based Rationalizations:** For some actions, A’s action is done for a reason if and only if (1) A has some element in her subjective motivational set (usually desire) the satisfaction of which will be served by her φ-ing, (2) A has an emotion that represents a relation between herself and something she values in the world, and (3) A’s action falls within the normal range of action tendencies for the relevant emotion.

But so far, I’ve only been considering cases of expressive actions—a class of actions meant to be a challenge to the standard account. What about less expressive
cases? Imagine you’re on a safari vacation. You’re sitting around the campfire when you hear the piercing howl of hyenas. You and your fellow adventurers are very afraid. The hair is standing up on your arms and neck, your heart rate hastens, your palms sweat. As one hyena approaches you in particular, you puff your chest out and reach down for a log that’s only barely in the fire, prepared to swing its blazing end at the approaching animal. The animal cowers and runs away. What’s the best explanation of your action? A standard account would say that you desired to be alive, and you believed that to stay alive, you would have to threaten the hyenas with a fiery stick. But is that the best explanation? Is it even a full explanation? I think we ought to make reference to your fear in a complete explanation. But not only that, I believe that your fear can combine with your desire to stay alive to rationalize your action. And you’d probably say that too—“I was very afraid for my life.” You had no time or wherewithal to formulate any such belief as is required by the standard account. Nonetheless, what you did was perfectly reasonable.

In this chapter, I will argue that what seemed to be the best explanation of purely expressive actions in chapter 5 will also serve as an explanation of actions done out of emotions that are not purely expressive. I will first clarify my thesis as well as what it means to act for a reason. With such background in order, I will consider the case of Emily from Nomy Arpaly’s “Acting Rationally Against One’s Best Judgment.” Arpaly argues that Emily is acting for reasons according to the standard account, but I think the only way Emily could be acting for reasons is because the emotion is part of the reason for which she acts. Considering this example helps me
to defend in full the thesis that an agent can act for a reason by acting on an emotion. Then I consider a series of objections to this thesis and reject each of them. I conclude with some positive implications of the theory.

II. Acting for Reasons

The first stage of my argument is determining what it means, in general, to act “for a reason.” It is commonplace in philosophy to talk about “acting for a reason.” Almost everyone working in meta-ethics or philosophy of action engages in this talk. While disagreements on what a reason is such that one can act for a reason are familiar, there’s less discussion on the nature and force of the ‘for’. What is it to act for a reason? To act for a reason is to be responsive to that reason in acting. But what is the nature and force of this responsiveness?

I take Donald Davidson to be a canonical example of someone who gives an account of acting for reasons. According to Davidson, we explain an agent’s action by giving her reasons for doing what she did (Davidson: 1980, 3). Such reasons-explanations, or rationalizing explanations, are a species of causal explanation. What separates a reasons-explanation from a causal explanation is the rich notion of justification. Some causal explanations involve justification, and some don’t. Some actions are merely caused, and others are responses to justificatory reasons. Those that respond to justificatory reasons are rationalizing explanations. Rationalizations justify in the sense that there is something to be said for performing the action, from the agent’s point of view. We do not find such
justifications in mere causal explanations. I present below two broad views one
might have concerning acting for a reason. One view requires that an agent be
aware of the reasons for which she acts as reasons, whereas the other view requires
only that the agent be aware of her reasons. The former is slightly more demanding
of agents, so I shall call it the Inflated View, whereas the second view is less
demanding, so it shall be the Deflated view.

In very broad-brush strokes, an Inflated view requires one to have what I call
a “normative belief” to be considered as acting for justificatory reasons. A
normative belief is a belief with thus-and-so is a reason for me to φ as its content. So,
to have a rationalizing explanation for why one is running a mile, one must have the
normative belief with the content “that running is healthy, and I want be healthy is a
reason for me to go for a run.” Then, the agent must act because of this normative
belief. A Deflated view is less demanding in that it does not require these normative
beliefs. Rather, the non-normative belief that running is healthy, combined with a
desire to be healthy is all that’s required. We need not add the further claim that this
belief and desire together is a reason for me to go for a run. But the agent still must
act because of this non-normative belief and desire. So, the Inflated view and the
Deflated view differ on what sorts of beliefs are required for acting for a reason; but
they agree that to act for a reason, one must act because of one’s reasons, which are
psychological states. I call this the necessary ‘because’ relation.
Kieran Setiya is an example of someone who endorses the Inflated view (Setiya, 39-47). Setiya argues that his account implies a cognitive constraint on taking-as-one’s-reason such that, “in order to act because p, ..., one must take that consideration as one’s reason to act, and in doing so, one must believe that p. In other words, one must believe that one believes that p” (Setiya, 43). But believing that one believes that p (where p is a consideration that counts in favor of φ-ing) is not yet to have a normative belief as I’ve characterized it. Setiya argues, though, that believing that one believes that p is not a sufficient condition on acting for a reason. So, he adds in the belief that p gives one a reason to φ. The more accurate account of acting for a reason, for Setiya is:

To take p as one’s reason for doing φ is to have the desire-like belief that one is doing φ for the reason that p, not just that one is doing φ because of the belief that p (Setiya, italics mine, 45).

So, to act for a reason, one must have the attitude with the content \( p \) is a reason for me to \( \phi \). This attitude “must present itself as part of what motivates my action” (Setiya, 45). Without such a normative attitude, one is not acting for a reason. So, Setiya’s endorsement of what I’ve called the Inflated view is that:

This attitude [in acting because p, I take p to be a consideration belief in which motivates me to \( \phi \) because I so take it] does depict me as acting for a reason, since it depicts me as being motivated partly by itself, namely by the fact that I take p as my reason to act (Setiya, 45).

---

57 Kolodny expresses a commitment to the Conservative view with his Generalized Internalism Requirement: “p is a reason for A to R only if it is possible for A to reason from the content of the recognition that p is a reason for A to R to R-ing” Kolodny, Niko. “Why Be Rational?” p. 548
Without such a normative attitude, one cannot act for a reason.

I think the Inflated view demands too much from agents. Moreover, I bring up the Inflated view to show how much it demands, to ensure that these high demands are not placed on my own view of emotions rationalizing action. This Inflated view also demands far more than what is necessary for the standard account, and so it helps to clarify what the standard account is committed to. To see how over-demanding the Inflated view is, first consider the nature of human deliberation. When I’m wondering whether to order the fish in the reputable restaurant, for example, my thoughts inevitably turn to the facts. I ask myself: should I order the fish? I consider whether the restaurant is reputable. I consider the likelihood of the seafood being delivered fresh. I wonder if I’m more in the mood for chicken or steak. I don’t want to order something others are ordering. From these considerations, I decide to order the fish. I need not also add in the belief that since the restaurant is reputable, the fish is fresh, I am in the mood for fish, and no one else is ordering the fish all constitute a reason to order the fish. Rather, because of these reasons for ordering the fish, I simply order the fish. To first come to the conclusion that I have most reason to order the fish for my decision to be made for reasons at all would be redundant. When I decide to order the fish, I’ve decided because of my reasons. Any additional belief that I’ve decided for good reasons would not add to the deliberative process. The conclusion of the deliberation—an intention—just is deciding for good reasons. And I have directly responded to those reasons.
Second, consider young children who do not have the full concept of acting for a reason, and yet they do act for reasons. A newborn infant is significantly different from an eight-month-old. The newborn has very little control over his movements, and he relies primarily on instinct. But once the baby gains control over his body, he will show evidence of reason responsiveness. The eight-month-old is sitting in his high chair with some cheerios, some diced pears, and some soft carrots in front of him. He prefers the taste of the pears the most, and so he intentionally eats all of those first. Then he moves onto his second preference—the carrots. He leaves all the cheerios on his tray (or, more likely, throws them down to the floor).

We can imagine the eight-month-old’s thought processes. He wants the pears the most, and he has various beliefs about what they look like and where they are on his tray. Given that, he intentionally reaches out and grabs a piece of pear. What is not included in his thought process is anything about liking pears and believing the pears are there giving him a reason to pick them up first. Children—especially babies this young—simply do not have the concept of ‘this-is-a-reason-for-that’. Nonetheless, they act for reasons.

Of course, there are times when I do make the normative belief explicit in my deliberation. Sometimes, we have to make very difficult decisions. When I’m wondering whether to take a job in North Carolina or in California, I probably have good reasons for both. I also have reasons against both. I weigh these reasons up in a sort of “pro-con” list, and then say to myself “well, it looks like I have most reason
to take the job in North Carolina.” Then, given the normative belief, I decide to accept the job in North Carolina. My decision to take that job is responsive to my normative belief about the reasons I have. But these cases are few. We don’t make big decisions all the time. In normal, everyday deliberation, we form intentions because of the reasons we have without explicitly forming them for a normative belief that these are the reasons we have. Thus, while the Inflated view seems plausible for big decisions we make occasionally, it is vastly over-intellectualized for everyday decisions like running to the store or deciding what to order for dinner.

While the Inflated view demands the presence of a normative belief or attitude about the reasons one has, the Deflated view demands far less. As a canonical example of the Deflated view, Davidson argues, “to the relation between a reason and an action it explains is the idea that the agent performed the action because he had the reason” (Davidson: 1980, 9). For Davidson, “R is a primary reason why an agent performed the action A under the description d only if R consists of a pro attitude of the agent towards actions with a certain property, and a belief of the agent that A, under the description d, has that property” (Davidson: 1980, 5). Thus, we arrive at a representation of the Deflated view on which an agent forms an intention because she has a reason to. When she acts on that intention, she acts for a reason.

This ‘because’ relation is rather simple on the Inflated view—the agent just has a belief that such-and-such is a reason for me to φ. And then this belief motivates action. But things are a bit more tricky for the Deflated view. There are
the reasons, on the one hand, and then the action on the other. Then there’s “acting because of the reasons” that’s stuck in the middle. What could that be? Arpaly offers a thorough explanation of the ‘because’ relation between beliefs and desires and actions (or intentions) in “Reason Responsiveness in a Deterministic World.” Arpaly’s basic account of acting for a reason is that reasons are content-efficacious causes of action. However, to act for a reason, one needs to respond to it in virtue of its content. She states the view as such:

If the content of a group of mental states—under the most common story these would be beliefs and desires, but they do not have to be—adds up to compelling reasons to do something, and these mental states motivate me to do it by virtue of the very feature that makes them good reasons, good premises in a practical inference—then I have responded to a reason (Arpaly: 2006, 70).

A group of mental states might cause behavior in two ways. One way is mere content efficacy, and the other way is full-blown reason-responsiveness (Arpaly: 2006, 61-65). For both kinds of causation, one mental event causes another by virtue of its content. The crucial difference is how the two contents are related. In the case of acting for a reason, the content of the first mental event relates to the content of the resultant mental state in the way premises relate to conclusions in an inference. No such inference is present in mere content-efficacy.

---

58 Arpaly actually list three ways in which a behavior might be caused (Arpaly: 2006, 61-65). One way does not involve content at all. An example of this is the Opera singer hitting a high note, causing my ears to hurt, and thus causing me to cover my ears. Likewise, there might be abnormal brain activity in a person that causes an epileptic seizure. Neither of these examples involves a mental event causing some behavior, so neither is a candidate for reason-responsiveness. I am primarily concerned with the two different ways in which a group of mental states, not some external event, causes behavior. So, I won’t discuss this third way in detail.
For Arpaly, if a reason is something one can act for, then the agent must have some minimal awareness of the reason for which she acts. This might simply mean the agent has a belief with the reason as its content. Without a belief with the reason as its content, however, there would be nothing to reason from in a practical inference, that is, there would be nothing to motivate action. Without reasoning in the form of a practical inference, there would be no acting for reasons. The “because” relation simply ends up being an inference on the Deflated view.

This is Arpaly’s ‘in virtue of’ relation. The meaning of “in virtue of” in this context is nicely captured by her example of Emma:

For Emma to drop the book on George’s head in order to hurt him, it has to be true not only that the relevant belief and desire caused her to drop the book (Davidson’s condition), but also that the belief and the desire caused her to drop the book by virtue of the thing that makes them a [motivating] reason for her to do so. That is, not by other features they have, and not even by other features of their content, such as their unnerving nature, but by virtue of the fact that their contents makes them good premises for a practical inference that points toward dropping the book on George’s head (Arpaly: 2006, 71).

Certainly, the reasoning need not be conscious or deliberate. What matters is the connection between the content of some group of mental states and the action performed in virtue of that content. Arpaly suggests that the “right sort of content” for reason-responsiveness is the kind that would figure as a premise in a practical

---

59 This need not be a belief with “such-and-such is a reason for me to φ.” That would be the Inflated view. Rather, the reason to blow out the candle is that it might start a fire, and you don’t want a fire. All I’m saying here, with respect to the Deflated view, is that the agent be minimally aware of the belief that the lit candle might start a fire, and that the agent be minimally aware of a desire not to have the house catch on fire.
The “right sort of content” for reason responsiveness on the Deflated view need not be normative in the way required by the Inflationist.

As I’ve presented things, the Inflationist gets reason responsiveness by requiring one to have some normative attitude towards the reason—a belief that $p$ is a reason for me to $\phi$. The Deflationist denies such a normative attitude, but rather places the emphasis on different ways in which reasons might be related to actions. By differentiating between mere causal efficacy and true reason responsiveness, Arpaly—a Deflationist—is able to capture the notion of acting for a reason without appealing to such normative attitudes.

I think the Deflated view is an improvement on the Inflated view. And notice, the Deflated view is consistent with the standard account, as I’ve described it. When responding to reasons, the reasons one is responding to are typically belief/desire pairs, and when the agent acts intentionally, she is acting for such reasons. However, Arpaly also considers that while acting for reasons is required for action, there might also be other, “non-content factors,” that help to explain one’s action. A good example of a “non-content factor” is being under the influence of alcohol. So, you’re really drunk, and you get in a fight with someone at the bar. This guy has been making crude jokes about your wife, shooting down your favorite

---

60 She also aptly notes that what matters is the kind of content. The assumed “carrier” of the content is the classic belief/desire pair, but that is not essential for reason-responsiveness even on her view.

61 One might say the two are completely interchangeable, in practice. However, there are reasons to think they are not in principle interchangeable.

62 I use the example of alcohol here, but as we’ll see, Arpaly also thinks emotions play this “non-content factor” role in explaining irrational action.
basketball team, and threatening to fight you all night. The guy’s behavior seems to
give you pretty good reason to deck him, so you do. Now, you are reason-responsive
so far—you’re acting for good reasons. But suppose that you would have never
done that if you hadn’t been drunk—you tend to get angry more easily when you are
drunk. Here is a case where the “non-content factor” actually helped you act on
your good reasons for acting. Being drunk helped you build up the courage to do
what you had good reason to do.

But consider another case. You’re drunk again. But this time, you’ve just
been laid off from work. You’re quite distraught about this, and you are primed for a
fight. You notice a guy sitting at the corner of the bar, and you just walk up to him
(unprovoked) and punch him in the face. You’re drunk and angry, just like the
previous case. But this guy has done nothing wrong to you, and so you have no
reason to punch him. In this case, the non-content factors have prevented you from
acting for good reasons—reasons not to punch the guy. Arpaly claims that “the
[non-content factor] only makes one irrational if it prevents one from acting for
some good rationalizing reasons” (Arpaly: 2006, 76). And so, your action is
irrational.

In what follows, I’ll have two tasks. First, I’ll show that there is a problem
with Arpaly’s account of acting for a reason—she’s being inconsistent. As I’ve
presented the view here, Arpaly claims that to act for a reason requires awareness
of the reason, and the reason must serve as a good premise in a practical inference.
But in a different paper, Arpaly presents a series of examples of agents acting for
reasons, however these agents do not have awareness of the good reasons they have, by Arpaly’s own lights. So there is a tension in her view. I go on to argue that these agents do in fact act for reasons, but not according to the Deflated view. In order to show how these agents are acting for reasons—a plausible claim—we must reference their emotions. Second, I need to say whether emotions are reasons for which one can act or simply non-content factors that influence our actions. That is, do emotions simply help us act for good reasons? Or are they reasons themselves? I’ll argue that emotions (combined with desires) can be reasons that rationalize action.

III. A Problem with Arpaly’s View of Reason Responsiveness

In “Acting Rationally Against One’s Best Judgment,” Arpaly gives us an interesting example. I will now re-examine that earlier example in light of the account of reason-responsiveness and “non-content factors” discussed above. Consider the case of Emily:

“Emily’s best judgment has always told her that she should pursue a Ph.D. in chemistry. But as she proceeds through a graduate program, she starts feeling restless, sad, and ill motivated to stick to her studies. These feelings are triggered by a variety of factors which, let us suppose, are good reasons for her, given her beliefs and desires, not to be in the program. The kind of research that she is expected to do, for example, does not allow her to fully exercise her talents, she does not possess some of the talents that the program requires, and the people who seem most happy in the program are very different from her in their general preferences and character. All these factors she notices and registers, but they are also something that she ignores when she deliberates about the rightness of her choice of vocation: like most of us, she tends to find it hard, even threatening, to take leave of a long-held conviction and to admit to herself the evidence against it. But
every day she encounters the evidence again, her restlessness grows, her sense of dissatisfaction grows, and she finds it harder to motivate herself to study. Still, when she deliberates, she concludes that her feelings are senseless and groundless. One day, on an impulse, propelled exclusively by her feelings, she quits the program, calling herself lazy and irrational but also experiencing a (to her) inexplicable sense of relief. Years later, happily working elsewhere, she suddenly sees the reasons for her bad feelings of old, cites them as the reasons for her quitting, and regards as irrational, not her quitting, but rather the fact that she held on to her conviction that the program was right for her for as long as she did” (Arpaly: 2003, 49-50).

Notice that good reasons for her not to be in the PhD program trigger feelings of restlessness, sadness, and a lack of motivation in Emily. She ignores these good reasons when she deliberates. Nevertheless, she quits the program on an impulse, propelled exclusively by her feelings. Years later, she sees the reasons for her bad feelings and cites them as reasons for her quitting the program. According to Arpaly, “Emily…acts far more rationally in leaving the program than she would in staying in the program, not simply because she has good reasons for leaving the program, but also because she acts for those good reasons” (Arpaly: 2003, 504-05).

I agree that just because Emily deliberated irrationally, that does not make her action irrational. I also agree that acting for good reasons does not require any deliberation at all—it does not require that one believe one is acting for good reasons, as the Inflated view demands. That is, I agree with the thesis that it is sometimes more rational to act against one’s best judgment. However, the point of interest here is how does Emily act for a reason? In what way is Emily reason-responsive? This is a crucial question because in order for Arpaly to say Emily was rational in action, Emily must be acting for a good reason. We have to find an
answer to this question in order to meet the demands of Arpaly’s larger project. Unfortunately, as I hope to show below, I don’t think Emily is reason-responsive on Arpaly’s account of reason-responsiveness. I believe Emily is reason responsive nonetheless. So, I will later provide an alternative account of reason responsiveness that captures the sense in which Emily acts for a reason.

Arpaly argues that to act for a reason, the agent must have some minimal awareness of the reason for which she acts, and the reason must cause the agent’s action. She argues that Emily changes her mind “rationally” and “for good reasons...as a legitimate response to good evidence of which [she] is aware, even though [she] does not deliberate [her] way into [her] actions” (Arpaly: 2003, 510). This might simply mean the agent has a belief with the reason as its content, in-line with the Deflated view of reason-responsiveness. Furthermore, without a belief with the reason as its content, there would be nothing to reason from in a practical inference. Without reasoning in the form of a practical inference, there would be no reason-responsiveness. Does Emily have a full-blown reason to quit the program?

A very easy way for Arpaly to argue that Emily has a full-blown reason is to argue that without such a reason, there is no ground for rational evaluation.\(^\text{63}\) Arpaly thinks that to be rational, an agent must act for good reasons. The whole point of the example is to show that Emily is more rational in quitting the program

\(^{63}\) While this may sound obviously question begging, I think it is in fact the only option open to Arpaly. She argues that Emily didn’t deliberate when changing her mind, and that deliberation isn’t necessary for acting for a reason, but she never says what Emily is doing that is necessary for acting for reasons.
than she would be if she stuck with her best judgment to stay in the program. So, if
Emily is rational, then she must be acting for good reasons. But this response is too
easy. In fact, it begs the question. What I am trying to figure out is whether Emily is
reason-responsive or not, which can be pulled apart from the question of whether
Emily is rational or not. And I’m challenging the more basic idea of responding to
reasons: how is Emily acting for a reason at all? I will argue that Emily is not acting
for a reason by Arpaly’s own lights.

We should consider the difference between the present Emily (Emily at the
time she quit the PhD program) and the future Emily (years later when she sees the
reasons for quitting the program). We might say of the present Emily, as Arpaly
does, that she denies the good reasons she has to quit the program. She rejects
them, and then ignores them when forming her all things considered judgment to
stay in the program. Maybe she was mildly aware of a sense that things weren’t
going well, but she pushed that down whenever she really thought about it. She
went so far as to think she was irrational for quitting the program. Was Emily aware
of her reasons to quit the program when she stormed out of the department chair’s
office? By Arpaly’s own lights, it seems not.

Moreover, Arpaly says present Emily was “propelled exclusively by her
feelings,” and not the “good reasons” future Emily cites. So, the good reasons were
not the actual cause of her action. By Arpaly’s own lights, Emily cannot be acting for
a reason if that reason does not cause her action. Present Emily is unaware—at the
time of acting—of the good reasons future Emily cites, and those good reasons do
not cause her action. So, I conclude that Emily is not in fact acting for a reason when she storms out of the office, by Arpaly's own lights.

But I still think Emily was more rational in quitting the program than staying in it, and I think this because Emily did have good reasons to quit the program. So, how is Emily reason responsive? I want to suggest her feelings play an important rationalizing role. It is her feelings of sadness and frustration that represent the reasons she has to quit the program. So, I want to argue that when Emily acts on her feelings, she does so as a way of responding to reasons. Emily responds to her reasons by acting on her emotions.

For Arpaly, to be reason-responsive, one mental event must cause another by virtue of its content, and the move from one to the other must be a valid inference. The resultant mental state is reason-responsive only if the fact that there is a valid inference here is part of the reason for the move from the one mental state to the other. These conditions are not met in the case of Emily. We could simply conclude that Emily is not acting for a reason, and that her good reasons to leave the program do not rationalize her action. But such a conclusion is premature. There might be a different kind of rationalization—a different way of responding to reasons—that makes Emily's action perfectly intelligible. I'll present such a view of acting for a reason below.
IV. Emotions as Reasons

In this section, I’ll remind the reader of my account of acting for reasons that relies on emotions. Then I’ll consider three other ways of rationalizing actions from emotions (1) the reason to feel is a reason to act, (2) the emotion rationalizes the belief/desire pair, but the belief/desire pair rationalizes the action, and (2) the emotion only motivates action (as a “non-content factor”), but it doesn’t rationalize action. I’ll reject each of these alternatives as inadequate, which will serve to solidify my account as the most plausible one.

IV. a. Emotion-Based Rationalizations

In chapter 5, I argued for my view of expressive actions, the Emotion-Based Rationalization view. According to this view, an agent’s emotion can combine with her desire to rationalize action. Jane’s anger combines with her desire to scratch out Joan’s eyes in a photograph to rationalize her action. Not all actions can be rationalized by emotions this way—only those within the acceptable range of action tendencies for emotions. Jane’s action looks enough like retaliation—the typical action tendency for anger—to count as a rationalization. If Jane were, instead, to click her tongue twice and spin around out of anger, her action would not be so rationalized because it doesn’t look enough like retaliation. So, an expression of an emotion that is within the normal bounds of action tendencies combines with a goal-directed desire to rationalize action. This account will become more plausible as I distinguish it from three other accounts of actions done out of emotions.
IV. b. Emotion-Based Rationalizations Make Sense of Emily's Action

Given the Emotion-Based Rationalizations account, we can better explain the case of Emily. Emily's emotions correctly represent the relational properties of loss (sadness) and of having one's goals undermined (frustration). While her beliefs may miss this point—she continues to believe the program is a good fit—her emotions correctly represent the situation. Her sadness and frustration are fitting because she really is faring poorly in the program. These fitting emotions combine with a desire to storm out of the chair's office to rationalize Emily's action. The emotion here has taken the place of belief in a standard belief/desire pair. Instead, what we have is an emotion/desire pair. The expected belief, "this program is a bad fit for me" has been explicitly rejected by Emily, but her emotions which represent what truly matters to her are taking over. These rational emotions then combine with a desire in the appropriate way to make sense of the present Emily's actions.

IV. c. First Objection: Reasons to Feel are the Reasons to Act

There might be an alternate explanation, however. The future Emily, remember, later cites the reasons there were to feel sad and frustrated. And she cites those as her reasons to quit the program. Maybe the correct explanation of Emily's action is that the good reasons were reasons to feel sadness and frustration, and they were also her good reasons to act, where the emotion simply motivated her action. But

---

64 This is assuming the embodied appraisal theory of emotions from chapter 2.
the good reasons to feel were the reasons that rationalize her action. I claimed that
the emotions themselves combined with desires to rationalize actions. But this
alternative says that the reasons to feel the emotion, instead, also rationalize the
action—mediated by emotion. Which explanation should we accept? We could
accept (1) reasons A, B, and C are reasons to feel emotion E, and reasons A, B, and C
also rationalize φ-ing, where the emotion only motivates one to act for the good
reasons one already has. Or (2) emotion E correctly represents relational
properties X and Y, and this emotion combines with a desire to rationalize φ. In the
following sections, I’ll argue for (2) and against (1).

IV. d. Response

A reason to be in some emotional state might not always thereby be a reason to act.
Consider an example offered by Monika Betzler. Two colleagues, A and B, go up
for early tenure review; A is granted tenure but B is not. There may be a reason for
B to feel envy at A’s promotion to tenure, yet it would be prudent for him not to
express his envy, given that A will now be in a position to rule on B’s own tenure
(Betzler, 452). B’s reason for feeling envious is that A got something B thought he
deserved. He might act on this envy by being standoffish towards A, talking poorly
about A behind his back, or trying to undermine A’s progress in a number of other

65 Notice—only my explanation challenges the standard account.
66 Betzler, Monika, “Making Sense of Actions Expressing Emotions,” dialectica Vol. 61, N° 3
(2007), pp. 447–466
ways. However, he really has no reasons to do these things because it would affect his chances, now, of getting tenure in the future.

One might respond to this objection (as Betzler responds) by simply making the connection between reasons to feel and reasons to act defeasible. So, if one has a reason to feel, then the agent also has a reason to act on that emotion in the absence of opposing reasons not to act. B does have a reason to act on his envy of A, but other reasons— that so acting would affect his chances for tenure in the future—override the reason to act given by the reason to feel envious.

This defeasibility response, however, is unsatisfactory. I might have a reason to have some emotion without a corresponding reason to act even if I do not have other opposing reasons not to act. That is, the very fact that I have a reason to feel gives me no reason to act on that emotion.

For example, if I am Jewish living in Nazi Germany, I have all the reasons in the world to be angry with the Nazis for the injustices they’ve committed against many others and myself. However, I—a single mother with two young children—have absolutely no reason to fight the Nazis, fix the injustices being committed, or secretly plan an assassination of Hitler. These may be appropriate reasons for someone else. But I am a single mother just trying to protect my children—there is no way I could do any of these things to the Nazis. If we assume that “ought implies can,” then there is no sense in which I ought to (have a reason to) fight the Nazis. But I surely can feel intense anger towards the Nazis. And I should—they’re
committing immense atrocities. So, I have reasons to feel angry, but those same reasons do not give me a reason to act in any way.

To have a motivating reason to act implies I can act for that reason. I cannot do anything to fight the Nazis—all I can do is hide and hope I'm not killed. So, I have no reason to fight the Nazis. But I certainly have plenty of reasons to feel angry and outraged with the Nazis. Saying I have a reason to fight the Nazis is like saying I have a reason to be a few inches taller. I have no such reason because I cannot do anything about my height. I might have a reason to wear high heeled shoes to look taller, just like I might have a reason to support the Allies in their fight against the Nazis—but neither of these are direct reasons to fight the Nazis or be a few inches taller. I cannot do either of these things, so I do not have a reason to do these things.

One might argue that the reasons I have to feel anger do give me a reason to act in a way that supports the Allies. So, reasons to feel do give one reasons to act. But notice that what I just called an action—supporting the Allies—might only be an attitude I have. To act on that attitude would require things of me that I cannot do. So, it remains the case that a reason to feel might give me no reason to act—while it might give me reason to have some other attitude.

**IV. e. Second Objection: Smith’s Account**

Smith argued that we could save the standard account from Hursthouse’s challenge by supplementing the belief.desire pair with an emotion along the following lines. For Smith, a Humean of the sort that accepts a belief.desire model of action (what
I’ve been calling the “standard account”) can supplement his view with emotions so long as the explanation in terms of emotion takes for granted some Humean explanation. The way in which the emotion explanation takes a Humean explanation for granted is that the emotion makes sense of the belief/desire pair rather than making sense of the action itself. It’s once removed from the action. However, the belief/desire pair makes sense of the action itself. So, the emotion explains the belief/desire pair, but only belief/desire pairs can explain action, for Smith. Take Smith’s favored example of a grieving man rolling around in his dead wife’s clothes. For Smith, the man’s action is explained by the desire to roll around in her clothes combined with the belief that by doing this-and-that he is rolling around in them. But since this belief/desire pair seems “distinctly unsatisfying,” Smith claims we must also reference the man’s grief. But the grief is supposed to explain the belief/desire pair—it’s his grief that explains his desire to roll around in his dead wife’s clothes and also his belief about how to do that. The grief doesn’t explain his action, actually rolling around on the floor in her clothes.

IV. f. Response

I think Smith is wrong here. I believe the emotion directly rationalizes the expressive action; the emotion need not first explain the belief/desire pair, where it’s the belief/desire pair that rationalizes the action. Notice, the belief/desire pair by itself doesn’t rationalize the action at all. Hearing that a man desires to roll around in his wife’s clothes combined with a belief about how to go about it doesn’t
make the man’s action make any sense whatsoever. If rationalizations are required to make the agent’s action intelligible, then this sort of belief/desire pair isn’t doing its job at all. And notice, Smith agrees by calling such an explanation “distinctly unsatisfying.” The belief/desire adds nothing to the rationalization of the action. Maybe such beliefs and desires are in the causal stream of the action. But not all causes of action are actual rationalizations of action. If the belief/desire pair alone doesn’t rationalize the action, then why think something upstream can cause it to magically rationalize the action?

In this case, the belief/desire pair doesn’t make the action intelligible. What does make the action intelligible? That the man was grieving. Without referencing the man’s extreme sadness, his action makes no sense. So, it seems the emotion and not the belief/desire pair is doing all the rationalizing work. Smith may be right about one thing—sometimes emotions cause desires. The man’s grief may be the cause of his desire to roll around in his late wife’s clothes. He wouldn’t have the desire without the grief, after all. But the emotion is still the thing which rationalizes the action even if it causes a desire, which then causes action. The emotion must be referenced in a full rationalization.

But there’s a bigger problem for Smith’s solution. I don’t think such a schema—emotions cause desires, but desires rationalize—could generalize. First, even if such a schema could generalize, the emotion would still be part of the rationalization. The desire alone wouldn’t make his action make sense. You’d still have to include his grief in a full rationalization of his action. Second, sometimes the
emotion and the desire occur at the same time without a causal link between them. Emotions have specific action tendencies. While rolling around in a late loved one’s clothes might be at the “bizarre” end of the normal range, other actions are more towards the middle of the normal range. The more “bizarre” the action is, the more there is a need for making sense of the desire that’s partly causing it. So, it makes more sense to think of the emotion as causing the desire.

But consider less symbolic cases. If I’m hiking in the dense forest and come across a poisonous snake, I feel fear and a desire to run away at the same time. It’s almost as if feeling fear just is running away. Or the flight response is even part of the feeling of fear. Fear is such a powerful, physical force felt in our bodies. It would be very hard to distinguish it from desires, as the cause of desire. I think a better rationalization of running away is that I feared the snake, and desired to not be bitten. But the desire is a long-standing desire of mine, surely not caused by this one instance of fear. We understand such long-standing desires, which help distinguish the normal range of action tendencies from emotions. We all have long-standing desires to be alive (fear), to do the right thing (guilt), to be treated fairly (anger/jealousy/envy), to be well respected in a community (shame/pride), to achieve our goals (happiness), and so on. Most actions out of emotions combine specific emotions with such long-standing desires to rationalize the action, where it’s not the case that the emotion causes the desire. The more symbolic or “bizarre” the action is, I agree, the more likely we are to think of the emotion as explaining a desire. But for emotions further into the normal range, the actions are best
explained by an emotion combining with some long-standing desire. So, Smith’s view won’t generalize to all actions explained by emotions. I’m not even convinced it captures the more expressive cases because the belief/desire pair doesn’t rationalize the action at all.

IV. g. Third Objection: Emotions as Non-Content Factors

But perhaps Arpaly has a different view in mind. Instead of arguing that emotions are responding to reasons to feel and act, they are acting as “non-content factors” simply pushing the agent to act for her good reasons. Arpaly’s picture might be that Emily has a reason to act, and her emotions act as a “non-content factor” to help her act for those good reasons. So, the emotion is merely a “non-content factor” that helps to fight those factors that would prevent one from acting for good reasons, and in this way enhance the efficacy of one’s already present motivating reasons (Arpaly: 2006, 76). On this view, emotions don’t rationalize at all. They are merely part of the causal story of action.

Arpaly might re-describe Emily’s action in these terms. Emily’s feelings of restlessness and sadness help to fight her considered best judgment from leading her to stay in the program. These feelings help Emily act on her motivating reasons to quit the program, thus contributing to her rationality. So, Emily has reasons to quit the program, and while her best judgment is hindering her ability to act for those reasons, her emotions are helping her act for them. But the emotions
themselves do not rationalize her action—they merely add a motivational “push” to get her to act for her good reasons.

IV. h. Response

However, I don’t think Arpaly has left herself this sort of explanation. Consider Alice, another example from Arpaly. Alice has the same good reasons that Emily has for leaving the program. The program is a bad fit for Alice. Also, Alice feels restless and sad. Alice, too, quits the program on a whim, propelled exclusively by her feelings, despite having judged that she ought to stay in the program. Given just this much, one would think Alice is no more or less rational than Emily. Alice could act for those good reasons because her “non-content factor” emotion helps her act for her good reasons for quitting the program. But Arpaly claims that Alice quits the program for bad reasons, and so Alice is irrational. Why are Alice’s reasons for quitting the program bad when they are the exact same reasons Emily has? Arpaly argues that Alice’s feelings of restlessness and sadness are responsive to the wrong considerations. Her feelings arise from low self-esteem rather than any real incompatibility between her and the program. That is, while she has good reasons to leave the program, Alice acts for bad reasons because her feelings of sadness and frustration are not caused by those good reasons. So, Arpaly concludes, Alice leaves the program for bad reasons because she feels restless and sad for bad reasons. Thus, the emotion is not meant to play a purely “non-content factor” role by Arpaly’s
own lights. Instead, in the case of Alice, the emotion ought to be responding to good reasons to feel for those good reasons to also be reasons to act.

So, it seems Arpaly must resort to the first, problematic way to understand responding to reasons by acting on an emotion: having a reason to have the emotion is therefore a reason to act on the emotion. But, as I showed, there are problems with this way of understanding the relation between reason-responsiveness and acting on an emotion. After ruling out these three alternatives to my view, I conclude that the best explanation of Emily’s action is that her feelings of sadness and frustration combine with a desire to quit. When she acts because of her emotion and her desire, she is acting for a reason.67

V. Emotions are Reasons

It makes more sense, now, to say that Emily quit the PhD program because she felt sad and restless. Furthermore, her quitting the program was rational because she responded to good reasons—where the good reason was her emotion. Arpaly grants that particular mental states—beliefs—give you good reasons to feel and act, why not think that particular mental states, namely emotions, also give you reason to act?

Consider another example. Suppose Doris simply lost track of time, and she forgot to pick her 10-year-old daughter, Josie, up from school. By the time Doris

67 Notice that Alice’s action is also rationalized by her emotions and her desire. Even though her emotions are not caused by her good reasons, her emotion is nonetheless fitting. So her emotion can combine with the desire to rationalize her action.
looks up, it’s already 7 o’clock, and she was supposed to be at the school by 3:30. A teacher was waiting with Josie all that time, trying to call Doris, but Doris had her phone off and didn’t receive the calls. When Doris realizes what she’s done, she sprints to the car, drives as fast as she can to the school, and runs up the front steps. When she sees poor little Josie sitting there, she is overcome with relief. She immediately gives Josie a big hug and kiss. Once the relief fades, Doris becomes distraught with guilt. She says over and over again how sorry she is for leaving her there. Then she moves on to apologize profusely to the kind teacher who waited there with Josie. She even offers to take the teacher out to dinner with Josie and her, but the teacher declines. Doris probably says, “I’m so sorry” a hundred times to the teacher and Josie by the time the whole ordeal is over. She sincerely promises that this will never happen again.

This example of Doris is supposed to be a completely normal, recognizable example of how emotions rationalize many actions in our day-to-day lives. Emotions do not only rationalize symbolic or expressive “bizarre” actions such as Jane’s; they also do not only rationalize actions where the agent is denying or unaware of the good reasons she has to act, as in the case of Emily. Doris’ actions seem perfectly normal to us, and she is very aware of the reasons she has to rush out the door and to apologize to Josie and her teacher. Nonetheless, it is Doris’ emotions that rationalize her actions. Doris’ fear rationalized rushing out the door, driving a bit faster than she normally would, and running up the steps. She was afraid something bad had happened to Josie. And then her overwhelming relief
rationalized her holding Josie so tight in her arms—she was relieved to see Josie all in one piece. And finally, her guilt rationalized her apologies and her offer to rectify the situation with a dinner for the teacher. The reason Doris acted in these ways was because of the emotions she was feeling. The reason she drove so fast was her fear; the reason she hugged so tight was her relief; the reason she apologized so many times was her guilt. Doris feels bad, and she wants to make it better.

But does the emotion rationalize better than a more typical belief/desire pair? Suppose we take the emotion out altogether in our story. Doris again realizes she’s late picking up her daughter. She believes she should’ve picked up her daughter hours ago, and she desires to pick her daughter up on time. And she believes that she has to leave now, drive a bit faster, and get into the building as soon as possible to pick up her daughter. Upon seeing her daughter, she desires to hug and kiss her, and she believes that by doing this and that with her body, she is hugging and kissing her daughter. She desires to make her daughter feel better, and so she apologizes for leaving her there. Lastly, she desires to make amends with the teacher, and believes that by inviting the teacher to dinner she is doing just that. Such a description leaves out reference to emotion. Does this story seem to make more sense than my previous, emotional story? If you’re a parent, which story seems more compelling? Which seems more true to reality?

I think most people would agree that the emotional story makes more sense—how could any parent not be feeling fear, followed by relief, followed by guilt? There’s something weird, or even “psychopathic” about the story that leaves...
out any reference to the emotions. In order for the story to make any sense at all, we must include Doris’ emotions, and whenever some state must be included for an action to make sense, then that state contributes to the rationalization of the action.

Emotions should count as an adequate, competing rationalization. They do not exhaust all possible rationalizations, of course. I simply want to show how they might be one possible rationalization amongst many. Given the fact that emotions represent relational properties that matter to us, what reason is there to think they cannot play a part in rationalizing action? I think the burden of proof is on my opponent here. So much of our lives are emotional. We regularly attempt to rationalize our actions by referencing our emotions. I slammed the door because I was angry. I pulled the plug on your project because I was envious of your success. I made us leave the party because I was jealous. We routinely cite our emotions as reasons for why we do the things we do. Sure, we might be able to concoct belief/desire pairs in each of these situations. But I don’t see any principled reason for ruling out the more natural emotion explanations just because we can concoct other explanations.

VI. Are Emotions Good Premises in a Practical Inference?

But there’s a lingering worry for my account. Arpaly’s account of reason responsiveness required good reasons to be the kinds of things that could serve as a premise in a practical inference, and that making such an inference amounts to acting for a reason. I want to argue, first, that one can act for a reason without
reasoning, and second, that something can be a good reason without being the kind of thing that can serve as a premise.

**VI. a. Reasoning**

We cannot say that for all reasons, if the reason justifies action, then the reason must’ve been the result of a piece of practical reasoning. Consider intentions. Intentional action is presumed to be action done for a reason, but it need not always be the result of *reasoning*. Intentions need not follow from some piece of reasoning in order to still be intentions. We might simply form them. And acting on those intentions that do not follow from a piece of practical reasoning still counts as intentional action—action done for a reason. For example, suppose Robert is sitting on the couch, and he suddenly says to himself, “you know what? I’m going to get up as soon as this show is over and wash the dishes.” While such an intention is reasonable—there are dirty dishes in the sink and Robert doesn't have anywhere else to be—he does not reason *from* the fact that there are dishes in the sink and he doesn’t have anywhere to be *to* forming the intention. It just sort of pops in his head. But now that he’s formed the intention, there is some reason to hold to it. If the show ends, and he just keeps sitting on the couch, then he’s done something less than fully rational.

Just because an intention does not follow from reasoning, that does not mean there is no reason to act on the intention. Likewise, just because an emotion does not follow from a piece of reasoning, that does not mean there is no reason to act on
that emotion. In this sense, emotions are like intentions. They need not follow from some piece of reasoning to be emotions. And acting on these emotions might be acting for a reason even if there was no reasoning involved. Feeling disgusted by the feces-shaped fudge is a reason not to eat it, even though there was no reasoning process leading to the feeling of disgust.

**VI. b. Premises**

But Arpaly argues that another essential feature that makes some mental states good reasons is that their contents form good premises in a practical inference. To show that emotions can be good reasons, I would have to either show that the content of emotions can form good premises in a practical inference or that Arpaly's condition on being a good reason is too strong. I'll opt to show the latter: something can be a good reason without being a good premise in a practical inference.

Requiring a good reason to be a good premise in a practical inference is too strong. The principle won't generalize to all good reasons. Consider perceptual experiences. Perceptual experiences often are good reasons to form perceptual beliefs, but the content of the perception is not something that could be a premise in a practical inference. Many philosophers of mind (Peacocke, Tye, Burge from chapter 3) argue—successfully, I think—that the representational content of a perception is nonconceptual. It is not structured in the way premises in a practical inference are.

---

68 This position is argued for more extensively in chapter 3. We can be entitled to believe that p on the basis of a perceptual experience of p even though p doesn’t serve as a premise in an inference to belief that p.
inference are. Yet, we rely heavily on these perceptual representations to justify our perceptual beliefs. So, there must be some sense of justification that is non-inferential. If there is a sense of justification that is non-inferential, then we can apply that kind of justification to the case of emotions. I'll return to such a sense of justification as discussed in chapter 3.

Tyler Burge views epistemic warrant as a genus with two sub-species. First, there is the epistemically internalist sub-species of epistemic warrant—justification. Justification is warrant that is “conceptually accessible on reflection to the warranted individual” (Burge: 2003, 505). That is, justification involves an individual having positive reasons (and being able to access those reasons) for what she believes. This is the sense of justification Arpaly had in mind. But there is another sense of warrant. Second, there is the epistemically externalist sub-species of epistemic warrant—entitlement. Entitlement is warrant that “need not be fully conceptually accessible, even on reflection, to the warranted individual” (Burge: 2003, 504). That is, a warranted individual need not have positive reasons for what she believes to be entitled to that belief. Burge’s surprising and enticing claim is, then, that “we are entitled to rely, other things being equal, on perception, memory, deductive and inductive reasoning, and on…the word of others” (Burge: 1993, 458). I then added emotions to the list of things we are entitled to rely on. I won’t re-hash all of those reasons here (I refer the reader to chapter 3, sections III and IV). But—taking those arguments for granted—we can act for reasons by acting for emotions even though such emotions are not premises in a practical inference. Instead,
emotions are reasons we are entitled to rely on in acting. But what makes some emotions good reasons while other emotional experiences are bad reasons? Again, return to perception.

What makes a perceptual representation a good justification for a belief is that the representation is veridical. Such a veridical perception is then a reason for believing on the basis of that perception. Thus, the feature that makes these perceptual states good reasons to form a perceptual belief is that their contents are veridical, i.e., they match the way the world is. It would be far too strong to argue that the feature that makes a perceptual state a good reason to form a perceptual belief is that its content forms a premise in an inference. Perceptual experiences are not of the same form as perceptual beliefs—they cannot play such an inferential role. And yet, they provide excellent reasons to believe.

Emotions function in the same way. Emotions can be fitting. An emotion is fitting if the relational properties it represents are really there—just as a perceptual representation is veridical if the world it represents is THE world. If the veridicality of a perceptual state makes that state a good reason for believing that p, then presumably the fittingness or appropriateness of an emotion makes it a good reason for believing, intending, or acting. So, while emotions may not play an inferential role, that does not mean they also cannot play a justificatory role. They justify in the same way perceptual experiences justify. We are entitled to rely on our emotions

---

69 Even a non-veridical perception can be a reason to believe, albeit maybe not a *good* reason to believe.
even though such emotions are not premises in a practical inference. And such emotions can be good reasons to act if they are fitting. So, Arpaly's requirements seem to be too strong.

VII. Concluding Remarks

Given the arguments above, then we can conclude that sometimes an agent’s action can be rationalized by an emotion/desire pair rather than by a belief/desire pair. Things can go wrong, for sure. Joe might be really angry with Sam, and this anger combines with a desire to kill Sam, which then rationalizes Joe’s action. I might be forced into saying that anger could rationalize killing someone. But the sense of ‘rationalize’ I’m using is a rather anemic sense. I don’t mean to use it as full-blown moral justification. But notice that many belief/desire pairs might also rationalize killing someone in this anemic sense of ‘rationalize’. Joe might really want to kill Sam, and he believes the best way to do this is to strangle him with piano wire. Presumably this belief/desire pair would rationalize Joe’s action if he were to strangle Sam with piano wire. So, yes, emotions might rationalize a wide range of evil and bizarre actions. But so do beliefs and desires. What matters is that the state doing the rationalizing is a reason. Beliefs, desires, and emotions are all reasons.

Emotions, beliefs, and desires often work together to adequately explain an action. For instance, I’m not going to be afraid of burglars if I don’t already have the belief that burglars steal your stuff and the desire to not have my stuff stolen. So, beliefs and desires might be upstream from the emotion. Moreover, the occurrent
fear won't overcome me unless I believe there is a burglar in my home or trying to get in. Suppose I hide under my blankets upon hearing the loud noise downstairs. I’m expressing my fear of the burglar in doing this, but I also desire to be out of sight and believe (probably falsely) that hiding under the blankets makes me invisible. Most of our actions are best explained—and even rationalized—by something more complex than a simple belief/desire pair. And somewhere in many of those rationalizations are emotions. That’s the crucial problem with the standard account. It vastly oversimplifies a very complicated story. The complicated—and correct—story must make reference to emotions in rationalizing human action.

The main thesis of this dissertation is that emotions are reasons for action. Chapter 2 presented an account of emotions that seemed phenomenologically accurate, and then I showed how we are entitled to rely on such emotions in action in chapter 3. Chapter 4 presented the standard account of action, and I argued that such an account is incapable of explaining expressive actions. In Chapter 5, I suggested we ought to add emotions to the list of mental states that can rationalize expressive actions. In this chapter, I've argued more extensively for the thesis that emotions combine with desires to rationalize many different actions—not only expressive ones. I hope that it is now clear that the standard account is inadequate. It cannot make sense of actions done for emotions.
Works Cited


References


