



University of Pennsylvania
ScholarlyCommons

Publicly Accessible Penn Dissertations

2018

Actually Embodied Emotions

Jordan Christopher Victor Taylor
University of Pennsylvania, jcvtaylor@gmail.com

Follow this and additional works at: <https://repository.upenn.edu/edissertations>

 Part of the [Philosophy of Science Commons](#), and the [Psychology Commons](#)

Recommended Citation

Taylor, Jordan Christopher Victor, "Actually Embodied Emotions" (2018). *Publicly Accessible Penn Dissertations*. 3192.
<https://repository.upenn.edu/edissertations/3192>

This paper is posted at ScholarlyCommons. <https://repository.upenn.edu/edissertations/3192>
For more information, please contact repository@pobox.upenn.edu.

Actually Embodied Emotions

Abstract

This dissertation offers a theory of emotion called the primitivist theory. Emotions are defined as bodily caused affective states. It derives key principles from William James's feeling theory of emotion, which states that emotions are felt experiences of bodily changes triggered by sensory stimuli (James, 1884; James, 1890). However, James's theory is commonly misinterpreted, leading to its dismissal by contemporary philosophers and psychologists. Chapter 1 therefore analyzes James's theory and compares it against contemporary treatments. It demonstrates that a rehabilitated Jamesian theory promises to benefit contemporary emotion research. Chapter 2 investigates James's legacy, as numerous alterations of his theory have influenced the field of emotion research over the past fifty years, including so-called neo-Jamesian theories. This chapter argues that all these variations of the Jamesian theory assume that emotions require mental causes, whether in the form of evaluative judgments or perceptual contents. But this demand is not present in James's theory. Nor, as Chapter 3 demonstrates, is this assumption necessary or even preferable for a fecund theory that explains human and non-human emotions. Thus, Chapter 3 offers the details of the primitivist theory of emotion: emotions are affective states that contribute to perceptual states by affectively representing relationships between the sensed environment and the sensing organism. Rather than relying on prior perceptual contents as triggers, emotions operate concurrently with, and as influencers of, exteroception. The information they carry can be conceptualized according to the theory of affordances proposed by ecological psychologist James J. Gibson (1979): emotions inform emoters of their potential responses to ecological concerns. Chapter 4 then explains how the primitivist theory is compatible with uniquely human emotion episodes, namely instances in which we identify and conceptualize our emotional experiences according to introspective and contextual cues. It argues that one resource available for categorizing our emotional episodes is a felt bodily map: different emotion concepts correspond with patterns of increased or decreased activity across the body. Finally, Chapter 5 situates the primitivist theory within the debate about natural kinds in psychology. While emotion constitutes a natural kind, discrete emotions do not.

Degree Type

Dissertation

Degree Name

Doctor of Philosophy (PhD)

Graduate Group

Philosophy

First Advisor

Gary Hatfield

Keywords

Affect, Consciousness, Embodiment, Emotion, Perception, Phenomenality

Subject Categories

Philosophy | Philosophy of Science | Psychology

ACTUALLY EMBODIED EMOTIONS

Jordan C. V. Taylor

A DISSERTATION

in

Philosophy

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2018

Supervisor of Dissertation

Gary Hatfield

Adam Seybert Professor in Moral and Intellectual Philosophy, University of Pennsylvania

Graduate Group Chairperson

Errol Lord

Associate Professor of Philosophy, University of Pennsylvania

Dissertation Committee

Karen Detlefsen, Professor of Philosophy and Education, University of Pennsylvania

Lisa Miracchi, Assistant Professor of Philosophy, University of Pennsylvania

Michael Weisberg, Professor and Chair of Philosophy, University of Pennsylvania

ACTUALLY EMBODIED EMOTIONS

COPYRIGHT

2018

Jordan Christopher Victor Taylor

This work is licensed under the
Creative Commons Attribution-
NonCommercial-ShareAlike 3.0
License

To view a copy of this license, visit

<https://creativecommons.org/licenses/by-nc-sa/3.0/us/>

For my family, old and new

Of four legs or two

Acknowledgments

There are far too many people deserving of my gratitude to name them all here, so I'll restrict my list of names to those who have directly and significantly impacted my thinking regarding this project.

Gary Hatfield, my advisor and my reason for applying to Penn in the first place, has been wonderful. Without his guidance and insights I would have been lost.

Each of my other committee members has also proven an invaluable source of wisdom. Thank you to Karen Detlefsen, Lisa Miracchi, and Michael Weisberg for keeping me challenged and motivated through your comments and discussions.

Deena Skolnick Weisberg and Emily Hopkins, my long-time colleagues in Penn's Cognition & Development Lab, helped me understand what goes on behind the scenes in experimental psychology. In collaborating with them I learned much about experimental designs and statistical analysis. Similarly, my knowledge of the concepts and methods of neuroscience was greatly enhanced by Martha Farah and Russell Epstein, two masters in their field who were instrumental in making neuroscience education accessible to students across the disciplines. Such knowledge enabled me to engage with scientific scholarship throughout my dissertation.

The philosophy department here at Penn boasts a very strong student population. I'm privileged to count myself among them. I must give particular thanks to Ben Baker, Devin Sanchez Curry, Louise Daoust, and Nabeel Hamid, my fellow members of the current

(2012–2018) iteration of the Hat Squad (a name first bestowed upon Gary Hatfield’s students by former member Larry Shapiro). Other students in the department who deserve my thanks for their consistent critical input include Marie Barnett, Justin Bernstein, Alkistis Elliott-Graves, Lindsey Fiorelli, Dylan Manson, and Brian Reese.

I’m fortunate to have had opportunities to present ideas and research from my dissertation at some excellent conferences and colloquia. At these events I’ve benefited from discussions with numerous scholars including Lisa Feldman Barrett, Randy Cornelius, Peter Godfrey-Smith, Heleen Pott, Jenefer Robinson, Andrea Scarantino, and Alex Skolnick. John Sutton and Paul Griffiths also deserves thanks for spurring my interest in the topic of emotion and the philosophy of science prior to my arrival at Penn.

Finally, thank you to my family, especially Devon for your unwavering support and enthusiasm. It’s tough work being a full-time academic; it must be tougher being married to one. And thank you to Toby for providing me with all sorts of anecdotal evidence and motivation to pursue this project. This dissertation is very much about how to understand you—but you’ll never read it, because you’re a dog.

ABSTRACT

ACTUALLY EMBODIED EMOTIONS

Jordan C. V. Taylor

Gary Hatfield

This dissertation offers a theory of emotion called the *primitivist theory*. Emotions are defined as bodily caused affective states. It derives key principles from William James's feeling theory of emotion, which states that emotions are felt experiences of bodily changes triggered by sensory stimuli (James, 1884, 1890). However, James's theory is commonly misinterpreted, leading to its dismissal by contemporary philosophers and psychologists. Chapter 1 therefore analyzes James's theory and compares it against contemporary treatments. It demonstrates that a rehabilitated Jamesian theory promises to benefit contemporary emotion research. Chapter 2 investigates James's legacy, as numerous alterations of his theory have influenced the field of emotion research over the past fifty years, including so-called neo-Jamesian theories. This chapter argues that all these variations of the Jamesian theory assume that emotions require mental causes, whether in the form of evaluative judgments or perceptual contents. But this demand is not present in James's theory. Nor, as Chapter 3 demonstrates, is this assumption necessary or even preferable for a fecund theory that explains human and non-human emotions. Thus, Chapter 3 offers the details of the primitivist theory of emotion: emotions are affective states that contribute to perceptual states by affectively representing relationships between the sensed environment and the sensing organism.

Rather than relying on prior perceptual contents as triggers, emotions operate concurrently with, and as influencers of, exteroception. The information they carry can be conceptualized according to the theory of affordances proposed by ecological psychologist James J. Gibson (1979): emotions inform emoters of their potential responses to ecological concerns. Chapter 4 then explains how the primitivist theory is compatible with uniquely human emotion episodes, namely instances in which we identify and conceptualize our emotional experiences according to introspective and contextual cues. It argues that one resource available for categorizing our emotional episodes is a felt bodily map: different emotion concepts correspond with patterns of increased or decreased activity across the body. Finally, Chapter 5 situates the primitivist theory within the debate about natural kinds in psychology. While *emotion* constitutes a natural kind, discrete *emotions* do not.

Contents

Preface	xii
Methodology: history and philosophy of science	xiv
Introduction	1
The primitivist theory of emotion	2
Theories of emotion from James to cognitivism	3
Overview of chapters	12
1 What is a Jamesian emotion?	18
Introduction	18
1.1 Standard interpretations and misinterpretations of James	22
1.1.1 James’s theory of emotion: key claims and hypotheses	22
1.1.2 Misinterpretations and confusions	24
1.1.2.1 James but not James-Lange	24
1.1.2.2 “Afraid because we run”: emotional and instinctual reac- tions	30
1.2 Are Jamesian emotions epiphenomenal?	35
1.2.1 An appraisalist interpretation	36
1.2.2 “Perception of the exciting fact”: sensation, perception, and cognition	41
1.2.3 Meeting the three intuitive desiderata	47

1.2.4	Micro-emotions and macro-emotions	51
	Conclusion	55
2	James's legacy: alternatives and alterations	58
	Introduction	58
2.1	Problems with cognitivism	60
2.1.1	Cognitivism is biologically narrow	61
2.1.2	Where do feelings fit in?	64
2.1.3	Emotion as a cluster concept	66
2.2	An unsuccessful union: cognition-arousal theory	68
2.2.1	Implications for the Jamesian program	70
2.2.2	Issues with Schachter and Singer's experiment	72
2.2.3	Issues with Schachter and Singer's theory	73
	2.2.3.1 The problem of metacognition	75
	2.2.3.2 The problem of the concept of self	76
2.2.4	How to feel yourself	78
2.3	What is a neo-Jamesian theory of emotion?	82
2.3.1	Antonio Damasio	82
2.3.2	Jesse Prinz	88
2.4	Do emotions require mental causes?	93
	Conclusion	96
3	A primitivist theory of emotion	97
	Introduction	97
3.1	Affect: its character, causes, and functions	99
3.1.1	The dimensions of affective qualities	100
3.1.2	Bodily causes of affect	102

3.1.2.1	Damasio and Carvalho's neurological account of feelings	103
3.1.2.2	Ginsburg and Jablonka's evolutionary account	106
3.1.3	Functions of affect	109
3.1.3.1	Affect's informational content	109
3.1.3.2	The intentionality of affect	110
3.1.3.3	Affect motivates action	111
3.1.4	Emotion is not merely affect	113
3.2	More than affect: emotion and perception	115
3.2.1	Affect and the evolution of perceptual systems	117
3.2.2	Affordances	120
3.2.3	Emotions and Ψ -affordances	125
	Conclusion	132
4	How to feel bad about yourself	133
	Introduction	133
4.1	The primitivist theory and its challenges	136
4.2	Accounting for diversity of phenomenality	141
4.2.1	Russell's three primitives	141
4.2.2	Bodily maps of emotional feelings	145
4.3	Dynamic episodes, static reports	147
4.3.1	Emotion concepts	150
4.3.1.1	Discrete concepts for gradient phenomena	153
4.3.2	Objects of emotion	158
4.4	How to feel bad about yourself	162
	Conclusion	165

5	Emotion is a natural kind, emotions are not	167
	Introduction	167
5.1	How to find a natural kind	170
5.1.1	Three approaches to natural kinds	171
5.1.2	Other kinds of kinds	175
5.2	Emotion is a natural kind	178
5.2.1	Arguments against emotion's natural kind status	179
5.2.2	Attempts to defend the unity thesis	181
5.2.2.1	A first step: affect program as an investigative kind	181
5.2.2.2	A second step: irruptive motivation as an investigative kind	183
5.2.3	Emotion's status under primitivism	185
5.2.3.1	Emotion is an investigative kind	186
5.2.3.2	Emotion is a natural kind	188
5.3	Emotions are not natural kinds	190
5.3.1	Some emotions may be investigative kinds	191
5.3.2	No emotions are natural kinds	195
	Conclusion	197
	Bibliography	198

Preface

When I think about how I landed at the idea for this dissertation, one memory from my first semester at Penn sticks in my mind. It was part-way through a seminar in Philosophy of Psychology: we were discussing the nature of cognition, or mental representation, or some similarly tricky concept. (Actually, maybe we were discussing *concepts*.) This being a philosophy class, no solid definition was forthcoming. Nor was much else agreed upon by everyone in the room. The assortment of senior undergraduates, graduate students at various stages of academic progression, and a professor well versed on the subject all muttered diffidently about it for a while. Eventually the professor, one Gary Hatfield, posed a question that brought the whole discussion into context. With a genuinely curious look on his face, he gazed out the window at the other end of the room and asked his students:

“When you put on pants, which leg do you put through first?”

As far as I was concerned, Professor Hatfield had just asked one of the most important questions in the philosophy and sciences of the mind.

Systematic inquiries into the mind tend to start at the top, so to speak. Philosophers ask lofty questions about the nature of the good or how one justifies one’s belief that Prince Charles has a belly button or how possibly one can *see* that-object-over-there *as* one’s copy of Homer’s *Iliad*. Neuroscientists use EEG and fMRI machines to scan the orbitofrontal cortex and the ventromedial prefrontal cortex (both are bits at the front of your brain) in search of the neural correlates of agency and decision-making. Cognitive psychologists look for

ever-finer distinctions between optimal and suboptimal conditions under which one can solve quadratic equations within an arbitrary time limit. (As it happens, having subjects sit on their hands diminishes their mathematical prowess.) Of course, no one should doubt that humans are remarkably intelligent creatures. Philosophers and scientists, as remarkably intelligent creatures themselves, want to explain the phenomena that make them and other humans remarkably intelligent.

But we seldom celebrate the putting on of pants (or other items of clothing) as a remarkably intelligent behavior. If a young professor writes a bunch of intelligent things about intelligent things, she is rewarded with tenure. If she turns up to work and proudly points out that she has succeeded in putting on pants, she is rewarded with strange looks and probative questions regarding her mental health.

While I've come to accept that no one will congratulate me for dressing myself, I'm disappointed that no one is really talking about it. More literally, I think it's a shame that philosophers and scientists of the mind tend to focus first on explaining those loftier human achievements rather than describing the everyday "thoughtless" behaviors that we consider unremarkable features of daily life. After all, just as our sophisticated linguistic capabilities and moral objectives and imperatives distinguish us humanfolk from other species, so too do more mundane behaviors such as dressing ourselves.

In fact, I'm most interested in explaining the psychological phenomena that we *do* share with non-human animals. Unfortunately that means putting aside explanations of how we dress ourselves. Such anthropocentric ambitions will have to wait until after I've dealt with more fundamental topics.

The topic of this dissertation is emotion. There's a bit in there about perception, too. These are two fundamental facets of mental life that we humans share with many other species. They precede or contribute to many of our most sophisticated behaviors. And so, I think these two phenomena should set the starting points for philosophical and scientific

inquiries into the mind.

This dissertation offers a *primitivist theory of emotion*, which sees emotion as one of the most basic aspects of consciousness. Emotions are bodily feelings that help to put us in touch with our environments. They contribute affective content to perceptual states and constitute a subjective, embodied *feeling-self*. Rather than understanding emotionality primarily as a sophisticated capacity exclusive to thoughtful creatures such as humans, primates, and some other mammals, the primitivist theory demonstrates that a massive range of Earth's creatures enjoy and suffer emotional experiences. It does so while avoiding an unsightly trade-off: though the primitivist theory expands beyond the scopes of competing theories the range of creatures whose mental lives involve emotionality, it does not thereby neglect the sophisticated experiences we humans celebrate and lament everyday. In this sense, not every primitivist emotion is a primitive emotion, so to speak.

Methodology: history and philosophy of science

The methodology employed throughout this dissertation integrates historical analyses, philosophical arguments, and reflections based on empirical investigations. A familiar label for this kind of approach is *history and philosophy of science*, or HPS for short.

While educating myself on the range of theories of emotion in the literature, I found that one of the most frequently-cited sources of inspiration (particularly in psychology) was a theory proposed by William James in the late 1800s. According to the current literature's understanding of this theory, an emotion is nothing more than a summation of felt bodily changes following the onset of some interesting stimulus. That is, an emotion is a feeling caused by the body. That seemed to me a good place to start an inquiry into the nature of primitive emotion.

There's a lot to unpack in James's theory. This is where the historical project of this disser-

tation comes in. Contemporary emotion theorists celebrate James as a pioneer but they no longer show much charity toward his theory of emotion. They understand it to be simplistic, empirically dubious, and superseded by theories informed by technological and conceptual advances in psychology, neuroscience, and philosophy of mind. This dissertation, however, offers a thorough analysis of James's theory. It shows that his major claims are still relevant to contemporary emotion research provided they are understood in the context of his broader psychological theory.

At the same time, all those progressions in sciences of the mind also demand our respect and attention. So, in its quest to revive a Jamesian theory of emotion, this dissertation discusses what that revival must entail given our up-to-date understandings of mind and behavior. But that doesn't mean conceding to every contrary claim published in a science journal. Good philosophy of science requires scrutiny of the science it aims to enlist for its explanatory purposes. I hope to have paid sufficiently close attention to the assumptions, implications, and biases of all sources cited in this dissertation regardless of their speculative or empirical natures.

James, by the way, is a remarkable figure in the history of philosophy of mind and psychology: he's often referred to as the "Father of American psychology". He was trained in medicine and philosophy, hired by his alma mater, Harvard University, to teach physiology and anatomy, and then he talked those in charge into letting him set up the first experimental psychology lab in the US. At one time he held positions in both the philosophy department and the psychology department at Harvard (the latter of which he established). He's my role model, you see.

While James worked across the disciplines, until very recently few contemporary academics can claim to have followed suit. This is certainly true of emotion research. Philosopher Louis Charland explains: "Although philosophical and psychological theorists of emotion do sometimes cite and refer to one another, they seldom engage in truly interdisciplinary

substantive analyses and discussion of common problems and concerns. As a result, there exist important issues that might benefit from an interdisciplinary dialogue that do not get raised or discussed in such terms” (Charland, 1997, 555-556).

Happily, the past two decades has seen significant progress towards interdisciplinary ventures into understanding emotion. Scientifically educated philosophers and philosophically educated scientists are realizing that they require one another’s expertise. And since the topic of emotion permeates numerous humanistic disciplines, such as political science, literary studies, and anthropology, one can make a case for conceptual and methodological cross-pollination not just between philosophy and the sciences of the mind, but between a broader spread of disciplines. Philosophy, as I see it, is a bridge between the sciences and the humanities; such is necessary for a plausible theory of emotion. In building the primitivist theory of emotion, this dissertation bolsters and reinforces this new interdisciplinary approach to emotion research.

Introduction

Emotions are very familiar, ancient features of our lives. But despite our intimate relationship with them, we struggle to understand exactly what they are. In 1884, William James famously asks: “What is an emotion?”—a simple question, but a tricky one. A full century later, psychologists Beverly Fehr and James Russell open their article with an astute observation: “Everyone knows what an emotion is, until asked to give a definition. Then, it seems, no one knows” (Fehr & Russell, 1984, 464). So much for scientific progress, you might think.

Human emotionality in particular is a slippery eel. Sometimes emotions can seem to be extraordinarily sophisticated responses to states of affairs that reflect our finely tuned social and moral norms. Moral philosophers write all sorts of books about them (e.g., Bell, 2013; Deonna et al., 2012; Martin, 2014; Nussbaum, 2001, 2004). Such projects are important because they can help us overcome or harness our emotions to the benefit of ourselves and others. Properly directed, emotions can even help us enact social and political change. As Rage Against The Machine vocalist Zack de la Rocha growls on the track “Freedom”: “Your anger is a gift”.

Sometimes, however, emotions can seem to be, or prompt, utterly stupid and inappropriate reactions to things that really aren’t significant to us at all. For example, in September of 2015 numerous local news websites in Indiana ran the following headline: “Woman jumps from moving car after seeing spider, causing wreck”. Across the Atlantic, in June of 2018, Wales News Service posted their own take: “Bear statue to be banned from Welsh town after

motorist thought it was real and crashed”. Wildly overblown emotional behaviors such as these can be very damaging, and not just to people’s egos. These two headlines alone justify a research project focusing on the nature of unsophisticated, primitive emotionality.

The primitivist theory of emotion

So, in examining the nature of emotion, this dissertation examines the minimal criteria for a mental event to qualify as an emotion. I’ve tried to find the simplest way to understand the term *emotion* and to determine what is at work in producing emotion in all manner of creatures. At the same time, I’ve ensured that the proposed theory—what’s called the *primitivist theory of emotion*—treats emotion as a functional phenomenon. In evolutionary terms, emotion obtains in Earth’s creatures because it helps us to survive and navigate our environments. (Except, as in the two headlines above, when it doesn’t.)

While the text shows a clear bias toward human emotional experiences, it also mentions numerous non-human animals: dogs, a bat or two, snakes, some birds (including an eagle and a self-knowing chicken), a mule, a bee, fruit flies, marine snails, porpoises, non-human primates, cats, mice, deer, sharks, jellyfish, an historically famous bear, and a made-up sea creature I call a “solipsistic bivalve” who may or may not be an ancestor of some extant species of mollusk. There’s also a rat, a rabbit, another dog, and a pigeon coming up in a few pages. According to the primitivist theory, it’s likely that most of these creatures (barring the solipsistic bivalve) enjoy a capacity for experiencing emotional feelings. And emotional feelings *are* emotions proper.

More precisely, an emotion is a particular kind of *affective state*. Affective states are sometimes called “raw feelings”. Whenever you feel hungry, or thirsty, or like you need to go to the bathroom, you are experiencing an affective state. Whenever you’re startled by a sudden change in your surroundings, or scared by some potential threat to your person, or thrilled

that something wonderful has just happened to you, you are experiencing an affective state. Orgasms, pains and aches, fatigue, cravings, and queasiness are all affective states. Affective states function to bring to an organism's attention the current status of its body. Hunger and thirst informs it that it's low on fuel. Tightness of the bladder and bowels informs it that it's time to release some waste. Fatigue informs it that it's time to get some rest.

As a particular kind of affective state, a primitivist emotion too informs the organism of the current status of its body. It also functions to relate the current state of the organism's surroundings to its body. Emotions are *about* not just the body, and not just the world, but the body–world relationship. This makes emotions their own kind of affective state, a kind irreducible to “mere affect”. (In fact, Chapter 5 argues that emotion qualifies as a natural kind in psychology.)

Furthermore, emotions partly constitute perceptual states. They operate concurrently with sensory registration, relaying to consciousness the needs, concerns, and capabilities of the body with respect to what is being sensed in the moment. In this way each emotion imports a self-interested affective quality to its perceptual states. We perceive objects by perceiving what we can do with or to them—or what we can do to distance ourselves from them or diminish their presence—and we gain this kind of information through our emotional experiences.

Theories of emotion from James to cognitivism

I should note that the primitivist theory wasn't conceived from nothing. Rather, it is the product of a research project that is roughly equal parts historical and systematic. The major claims of the primitivist theory derive from those espoused by William James, especially in his 1884 paper on emotion and his 1890 mammoth text, *The Principles of Psychology*. The general idea that emotions are bodily feelings, per primitivism, is the same claim for which

James is most well-known in the history of emotion research. (But, we will see, how to understand the nuances of that claim will differ markedly between James's theory and primitivism.) Primitivism also shares with James's theory the claim that emotions proper cannot be taxonomized into discrete psychological kinds such as *anger* or *fear* or *joy*.

For James, the conscious mind is a direct product of the dynamic body. Whenever an organism experiences a change in its mental state, this is necessarily due to some change in its bodily state. And the body is armed with an indefinitely large range of instinctual behavioral reactions to different stimuli; even more reactions can become habituated through association and training throughout the organism's life. James identifies emotions, or what he calls *emotional experiences*, as feelings caused by *emotional reactions*, those changes in bodily state that are triggered by *interesting* stimuli: features of the external environment (and in adult life, also features of one's concept of self) that promise to harm or benefit one's embodied, conscious, *nuclear self*. In fact, as Chapter 1 will demonstrate, James understands emotionality as central to both one's perceptual states and one's nuclear self.

Unfortunately James's theory of emotion is not paid much serious attention these days. A major reason for this is a series of tidal shifts in Anglo-American psychology since the early years of the twentieth century. During the Dark Age of Behaviorism (approximately the first half of the century), emotions often are conceived merely as patterns of behavioral and physiological responses to external stimuli; in extreme iterations, feelings are banished from the behaviorists' ontology. Take Carl G. Hempel's description of the general method of behaviorism:

Its principal methodological postulate is that a scientific psychology should limit itself to the study of the bodily behavior with which man and the animals respond to changes in their physical environment, and should proscribe as nonscientific any descriptive or explanatory step which makes use of terms from introspective or "understanding" psychology, such as 'feeling', 'lived expe-

rience,' idea,' will,' intention,' goal,' disposition,' representation.' (Hempel, 1980, 16)

John Watson's behaviorism champions the principle that all psychological explanation must be grounded in objective measurements and observations. Applying this to emotion research, the behaviorist "could simplify the problems of emotion and apply objective experimental methods to their solution" (Watson, 1930, 140).

That solution, as Watson sees it, is necessitated by the damage wrought by James's theory: "Nearly 40 years ago James gave to the psychology of the emotions a setback from which it has only recently begun to recover" (Watson, 1930, 140). James's theory—and his general approach to psychology—relies on *introspection*, the turning of one's attention inwards and analyzing one's own mental states. But as Watson's compatriot B. F. Skinner writes on behalf of behaviorism: "Introspective psychology no longer pretends to supply direct information about events which are the causal antecedents, rather than the mere accompaniments, of behavior. It defines its 'subjective' events in ways which strip them of any usefulness in causal analysis. The events appealed to in early mentalistic explanations of behavior have remained beyond the reach of observation" (Skinner, 1953, 30). James's theory of emotion, as a *psychological* theory, is problematic to the behaviorists because it relies on, and describes, subjective events that cannot be directly measured or objectively observed. Therefore, the Jamesian conception of emotion does not offer anything of use in causal analysis. That's the setback Watson sees. Feelings have no place in a respectable, "scientific" theory of emotion.

So what is the behaviorists' solution? To reconceptualize emotion according to what *can* be observed in human and animal behaviors. Watson developed an empirical project to examine and define emotions as patterns of behavioral responses to stimuli. You may be familiar with the story of Little Albert, a wee infant who was enlisted by Watson and his labmates to be traumatized in the name of respectable science. Watson gives a brief biography:

We chose as our first subject Albert B., an infant weighing twenty-one pounds, at eleven months of age. Albert was the son of one of the wet nurses in the Harriet Lane Hospital. He had lived his whole life in the hospital. He was a wonderfully “good” baby. In all the months we worked with him we never saw him cry until after our experiments were made! (Watson, 1930, 159)

Those experiments entailed handing Little Albert cute, fluffy animals to befriend (a white lab rat, a rabbit, a dog) before bashing a steel bar with a carpenter’s hammer right behind the boy’s head, and then seeing how many repetitions it would take to condition Albert to become apprehensive of cute, fluffy animals (Watson, 1930, 158-164). Understanding “emotion”, in this case, entails predicting and differentiating the patterns of behavior a subject produces in response to different kinds of stimuli: *stimulus–response pairings*, for short. And that’s that for the behaviorist: “It seems fair to call all of this group *emotional* without further defining the word at the present time” (Watson, 1930, 146).

In fact, Skinner was uneasy about using the term *response* to describe subject’s behaviors following stimulus onsets since the term carries some conceptual baggage. Labelling a behavior “a response” denotes that it is necessarily elicited by the stimulus (to which it is a response). But this assumption oversteps the objective observations of the scientifically respectable psychologist. Skinner writes:

It is customary to refer to any movement of the organism as a “response.” The word... implies an act which, so to speak, answers a prior event—the stimulus. But we may make an event contingent upon behavior without identifying, or being able to identify, a prior stimulus. We did not alter the environment of the pigeon to *elicit* the upward movement of the head. It is probably impossible to show that any single stimulus invariably precedes this movement. Behavior of this sort may come under the control of stimuli, but the relation is not that of elicitation. The term “response” is therefore not wholly appropriate but is so

well established that we shall use it... (Skinner, 1953, 64, original emphasis).

It makes sense that hard-line behaviorists like Skinner would prefer a different term. After all, their principles discourage invoking any concepts pertaining to agency, as evidenced by the quotation from Hempel a few pages above. Rather, their psychology explains the behaviors of organisms much like physics explains the “behaviors” of physical substances. A physicist wouldn’t describe an eight-ball’s movement toward the corner pocket as a “response” to the cue-ball’s hitting it.

Later into the twentieth century it became apparent that despite their noble ambition of physicalizing psychological explanation, the behaviorists had instead given the sciences of the mind another setback. To quote an audience member at a cognitive science conference I attended a few years ago: “There was nothing wrong with behaviorism except that it was incomplete”. What was missing was an account of the fact that our behaviors truly *are* responses to the world. Psychologists cannot ignore that fact and claim to make real progress in understanding the human psyche.¹ But in order to respond to something, subjects arguably require a capability to interpret that something as *something to be responded to*. Researchers started to realize that there must be some sort of information processing involved between stimulus onset and manifestation of behavioral responses, and that accounting for those processes is integral to explanations of their phenomena.

The so-called cognitive revolution promised to rescue internal psychological mechanisms from the behaviorists’ abyss by describing psychological events (*qua* responses) in computational terms. Its seeds were sewn at the end of World War II with certain developments in computational science. Psychological explanations saw stimulus–response pairings replaced with stimulus–*process*–response sequences. Regarding emotion, psychologists be-

¹A small subset of moderate behaviorists agree with this point and thereby reject the more extreme positions advocated by the likes of Hempel, Watson, and Skinner. Edwin Holt, for instance, pushes a behaviorism that involves notions of agency and intention (e.g., Holt, 1914, 1931). His student, Edward C. Tollman, developed a position called *purposive behaviorism* whose explanations include descriptions of the purposes and goals of the behaviors examined (Tollman, 1932).

came interested not just in determining the range of stimuli that will trigger a given pattern of behavior, but also in describing the internal processes by which subjects *interpret* that range of stimuli relative to their behavioral responses.

This *cognitivist* approach sees theorists prioritize information processing in their accounts of emotion. Today, cognitivist theories dominate the field of emotion research. Naturally, being so popular as a general approach to emotion research, cognitivism comes in a few different flavors.

In American psychology since about 1960 a mainstream approach is to postulate that emotions—whatever their nature—must be responses to *evaluative appraisals* of their objects (e.g., Arnold, 1960a,b; Lazarus, 1982, 1991, 2001; Scherer, 1993, 1999). Internal discussions among appraisal theorists focus on the means by which emotions are differentiated according to their antecedent appraisals, and how best to analyze what is represented by an evaluative appraisal when it triggers one emotion rather than any other (Scherer, 1999). If you experience fear after seeing a spider, it's probably because you appraise that spider (especially its venom-injecting fangs) as *dangerous*, a *threat to your well-being*. If you experience anger after someone insults you, it's probably because you appraise that person (or her utterance) as *offensive*, a *threat to your reputation and dignity*. Since this approach prioritizes inquiries into the causes and origins of emotions, Andrea Scarantino (2010) calls this species of cognitivism *Etiological Cognitivism*.

A corresponding movement in philosophy sees the revival of a Stoic idea: emotions are *constituted by* evaluative judgments. Martha Nussbaum explains:

This view holds that emotions are appraisals or value judgments, which ascribe to things and persons outside the person's own control great importance for that person's own flourishing. It thus contains three salient ideas: the idea of a *cognitive appraisal* or *evaluation*; the idea of *one's own flourishing* or *one's important goals and projects*; and the idea of the *salience of external objects as elements in*

one's own scheme of goals. Emotions typically combine these ideas with information about events in the world; they are our ways of registering how things are with respect to the external (i.e., uncontrolled) items that we view as salient to our well-being. (Nussbaum, 2001, 4, original emphasis)

This approach is different than Etiological Cognitivism since emotions “are not responses to what happens but evaluations of what happens. And they are not responses to those evaluative judgments but rather they *are* those judgments” (Solomon, 1976, 127, original emphasis). Instead of being triggered by the evaluative outputs of appraisal mechanisms, emotions themselves *are* those evaluations. Scarantino calls this approach *Constitutive Cognitivism*.

Constitutive Cognitivists typically emphasize more than do Etiological Cognitivists the centrality of the *self* in a description of emotion. Robert Solomon, for instance, writes: “The ultimate object of our emotional judgments is always our own sense of personal dignity and self-esteem. Whatever its particular object and strategy, ...an emotion is ultimately concerned with personal status, self-respect, and one's place in his or her world” (Solomon, 1976, 129). Nussbaum echoes this point: “Emotions, I shall argue, involve judgments about important things, judgments in which, appraising an external object as salient for our own well-being, we acknowledge our own neediness and incompleteness before parts of the world that we do not fully control” (Nussbaum, 2001, 19). This is not to say that Etiological Cognitivists neglect the position of self, however. Richard Lazarus, a flag carrier for appraisal theory, stresses that “ego-identity...is involved in every emotion” (Lazarus, 1991, 100), where the concept of ego-identity “encompasses the *person-in-the-world*, which includes roles, relationships, and functions in society” (Lazarus, 1991, 101).

Overall, then, we can characterize a *cognitivist theory of emotion* as a theory in which emotions are defined either as responses to judgments (Etiological Cognitivism) or as judgments themselves (Constitutive Cognitivism), and according to which emotions are intimately related to a sophisticated, socially involved concept of self.

Although psychology was in better shape following the decline of behaviorism, it had not yet done justice to phenomenology. For a long stretch of the twentieth century, mainstream psychology and philosophy of emotion in the Anglosphere either neglected the felt aspects of emotion or tossed them aside as epiphenomenal by-products of otherwise functional processes. Moreover, theorists saw little reason to invoke the body in explanations of emotion. After all, if emotions (and other mental states) are explained as computations over information, physiological and phenomenological accounts are unnecessary and certainly insufficient: at best they are luxuries, at worst they are complications.²

Nowadays, thankfully, most theorists pay the body and its feelings more attention than they used to. But these considerations still don't take priority when it comes to defining emotion. For example, Solomon has shifted his opinion quite dramatically since he published his book in 1976. He now acknowledges that bodily feeling is "the 'missing' element in the cognitive theory of emotions" (Solomon, 2004, 88). While still maintaining that emotions are judgments, he makes clear that his understanding of the term *judgment* accommodates phenomenological concerns: "the judgments that I claim are constitutive of emotion may be nonpropositional and bodily as well as propositional and articulate. They manifest themselves as feelings" (Solomon, 2004, 88). For the later, affectively enlightened Solomon, emotional feelings are important, but only because they are the means for keeping his claim that emotions are judgments on the table following the criticisms of anti-cognitivist theorists (cf. Scarantino, 2010). Nussbaum (2001) similarly conceptualizes the "feelings" of emotions as a species of judgment. On the other hand, she also states that any felt aspect of emotionality is a contingency due to the fact that we happen to be embodied minds: hypothetically, a disembodied mind could have an emotion absent any feeling at all.

In fact, feelings remain a big problem for cognitivist theories, as Chapter 2 demonstrates.

²This general attitude towards physiology's uselessness to psychological explanation is evident in Jerry Fodor's influential paper "Methodological solipsism considered as a research strategy in cognitive psychology" (Fodor, 1980).

Cognitivists of either camp don't usually try to banish feelings from their theories as the behaviorists do, but they struggle to find a functional role for them in their computational or propositional explanations. While progress in psychology and philosophy of mind since behaviorism has led to the resuscitation of mental states as internal computations, the phenomenological side of the mind waits impatiently for its turn under the spotlight.

The primitivist theory advocated here is a direct response to cognitivist theories of emotion. It revives the Jamesian claim that emotions are bodily feelings without thereby rendering emotions epiphenomenal. It maintains, against Etiological Cognitivism, that an emotion does not require an antecedent mental state (such as an evaluative appraisal or an informationally rich perceptual state) as its cause. It also holds, against Constitutive Cognitivism, that emotions are not judgments even though they are intentional by their nature, functioning to relate an organism to its world. What emotions do require as their cause is a body complete with interoceptors (intrabodily monitoring systems) and exteroceptors (sensory systems sensitive to the extrabodily environment). And while, as the cognitivists maintain, they are intimately related to one's self, that self is manifested by the body and its feelings—not contingently, but necessarily. The primitivist theory is a theory of *actually embodied emotions*: emotions that originate in, function in service of, and reflect the needs, concerns, and capabilities of the body and the self that it comprises.

In this sense the primitivist theory is a *noncognitivist* theory: emotions are neither necessarily responses to judgments nor judgments themselves. Importantly, however, the primitivist theory is not an *anticognitivist* theory of emotion. While a noncognitivist theory denies the necessity of antecedent mental states, it nonetheless allows that emotions may be triggered by events like perceptual and cognitive representations. We humans often produce emotional reactions when imagining being insulted, or remembering that we left the stove on, or upon learning that our friends are expecting a child. An anticognitivist theory, contrarily, denies that any such cognitive events can trigger an emotional reaction. The be-

haviorist view outlined above qualifies as an anticognitivist theory. (But note the difference between an *anticognitivist theory* and an *anti-cognitivist theorist*, a term I used in the last paragraph and one that Scarantino (2010) uses also. This latter term refers to an emotion theorist who opposes cognitivist theories. One can be an anti-cognitivist theorist without endorsing an anticognitivist theory!)

Nor does the primitivist theory deny that emotions contribute to cognitive processes. In fact many theorists will shudder at the quite narrow sense of the word *cognitivist* as it applies to cognitivist theories of emotion. The fact of the matter is that *cognition* is a bit like *emotion*: no two theorists agree precisely on what it is. Chapter 1 (subsection 1.2.2) demonstrates the danger of misinterpreting another theorist's understanding of the term. Throughout the dissertation you'll see quotations regarding cognition from theorists who bolster arguments in favor of the primitivist theory (especially from the works of embodied cognition theorists). I invite those with more liberal understandings of cognition to interpret *cognitivist theory of emotion* as a conventional term. When I use the term *cognition* in the text, I use it in the conventional sense per cognitivism about emotions.

Overview of chapters

Chapter 1 is devoted to detailing and analyzing James's theory of emotion. It introduces the theory's major premises and describes their components according to the broader psychological theory James developed throughout his *Principles of Psychology* (James, 1890). The objective is to demonstrate that James's theory of emotion, properly understood, offers much of theoretical value to contemporary emotion research.

The chapter compares this contextualized reading against two common misreadings found in extant commentaries. First, it disentangles James's theory from the James-Lange theory. The James-Lange theory is an ugly hybrid of James's and Carl Lange's superficially

similar but independently developed theories of emotion. There are a few reasons to keep these two separate. As well as disagreeing over how to define emotion (James says they're feelings; Lange says they're patterns of physiological activity), they differ in their accounts of the physiology underlying emotion. Moreover, while Lange's theory drifts very close to a behaviorist account, James's emphasizes the centrality of the self in conscious experiences such as emotions. The second misreading attributes to James's theory full-blown, world-directed behaviors (such as running away from a bear) as necessary causes for emotions. This is a strawman fallacy, rendering James's theory wildly implausible when in fact he means something much subtler by the "stirrings of the body" that trigger emotional experiences.

The second part of Chapter 1 investigates whether Jamesian emotions perform some sort of psychological function. Contemporary cognitivists demand that emotions perform a functional role in the mind; a charitable, cognitivist reading teases out of James's theory an appraisal mechanism whose operation helps to ground emotion's intentional character. A closer look at James's writings not just on emotion, but on perception and cognition, shows that this cognitivist interpretation is inaccurate. In fact, Jamesian emotions are noncognitive since they do not require prior mental states as triggers. But they are intentional: they inform the organism of the current state of its body and how that state relates to stimuli registered by the senses. In this way, Jamesian emotions import an affective, intentional character to perceptual states.

Chapter 2 explores some of the theories that have arisen since the decline of behaviorism. It begins by discussing the tenets of cognitivism and three prominent challenges cognitivists struggle to overcome. One problem is that while cognitivism may work well enough for describing human emotions, it fails to accommodate the emotions of many other species. Another problem is accounting for emotional feelings in cognitivist accounts, as mentioned above. And a third problem regards how to define an emotion: is it a single thing—a judgment, perhaps—or is it a number of different events (perhaps physiological, probably affec-

tive, definitely cognitive) grouped together somehow? If the latter, then what makes that collection *an emotion*?

After raising these issues for cognitivism, the chapter turns to an influential attempt to marry cognitivism with the Jamesian theory. According to this *cognition-arousal theory*, an emotion is a function of vague bodily feelings consciously interpreted according to a stock of available emotion concepts. But the cognition-arousal theory inherits from its cognitivist roots those very issues the cognitivists are yet to overcome. A major issue is that it demands that any organism properly capable of experiencing emotion must have an explicit concept of self to which it can attribute its own emotions during metacognition. This is an unnecessarily anthropocentric construal. Nor, as the chapter argues, is possessing a concept of self required for an emotional experience to be *about* one's self. In contrast with a *concept of self*, primitivist emotions realize and intend towards an embodied *feeling-self*.

The third section of Chapter 2 then looks at two more recent attempts to revive approximations of Jamesian theory in contemporary neuroscience and philosophy of mind. These so-called neo-Jamesian theories, as it turns out, are closer in spirit to the James-Lange postulate than James's actual theory. Both accounts see emotions as responses to prior mental states, namely perceptual states. But just as this assumption is challenged in Chapter 1 regarding James's theory, so too can it be challenged from a contemporary standpoint. Chapter 2 closes with a suggestion that a contemporary theory can allow for emotions that are logically prior to perceptual and cognitive representations.

Chapter 3 sets the stage for, and then introduces, the primitivist theory of emotion. This chapter is split into two sections. The first describes the concept of affect: how it is characterized across qualitative dimensions, how it is caused by the body, and how it functions for the conscious organism. While affective states carry information about the organism's bodily state and motivates the organism to act according to their characters, this alone is insufficient to explain the nature of emotional experiences (or emotional feelings). In the

second section, emotion is described as the product of a general affective system responding to activity of the body's exteroceptive systems. Emotion is therefore intimately connected to perception. To understand the relationship, this second section explores accounts of the evolution of perceptual systems, and discusses how a species's general affective system might have guided its perceptual development throughout generations.

In exploring the emotion–perception relationship, Chapter 3 makes a case for *affordances* as the primary contents of perceptual states. The notion of affordances originates in the works of James J. Gibson (1979). But the description of affordances offered in this chapter deviates somewhat from Gibsonian orthodoxy—depending, that is, on how one parses Gibson's own theory. The primitivist theory of emotion introduces a particular type of affordance, here named *Ψ-affordances*, that arise via an organism's emotional experiences while it perceives its environment. That is, emotions partly constitute the contents of perception by importing into them an affective quality reflective of the organism's bodily needs, concerns, and capabilities. This reverses the assumption, present in each of the theories explored in Chapter 2, that an emotion is necessarily a response to some prior mental state (such as a perceptual state or a cognitive state).

While the primitivist theory described in Chapter 3 can account for emotions in a huge spread of Earth's creatures, Chapter 4 narrows its focus to human emotionality. More specifically, it explores how human subjects come to identify and label their own emotion states. The chapter opens by detailing the challenges the primitivist theory must overcome in order to succeed as a comprehensive theory of human emotions. In doing so, it revisits and updates the cognition-arousal theory discussed in Chapter 2. It argues that typical human cases involve not merely introspectively analyzing one's emotional experiences, but also contextualizing them using a dynamic range of resources. That is, rather than attending to emotional experiences alone, we attribute emotions to ourselves by metacognizing over our *emotion episodes*. One resource for metacognition is the body: we learn to map out our emotional

experiences according to activity in different bodily regions. Another is using external cues that help to determine the onset and extinction of an emotion episode. A third is one's stock of emotion concepts, some of which may be embodied in nature. And a fourth resource comes in the form of objects of emotion: those features of the world that we attend to while we are experiencing emotions and that we interpret as significant to us in some way. These may be informed through ecological concerns (such as threats and benefits to one's bodily constitution) or social and moral norms (such as insults to one's dignity and deviations from ethical protocols). When emotions are *self-reflective* (such as guilt and shame and pride) they take one's concept of self as their objects. Overall, this chapter demonstrates that the cognition-arousal theory, with all its cognitivist promises, can succeed as a *theory of emotion self-reports*, but not as a theory of emotion *tout court*.

Chapter 5 then discusses two questions that are often confused as asking the same thing. First: Is emotion a natural kind? And second: Are discrete emotions natural kinds? The answers are yes and no respectively. The chapter opens by detailing what is meant by calling something a natural kind. There are three major positions on the matter: *conventionalism*, *essentialism*, and the *homeostatic property cluster* (HPC) view. Setting conventionalism aside, the chapter distinguishes between natural kinds as the kinds dealt with by essentialism, and their metaphysically looser counterparts, *investigative kinds*, as the kinds central to the HPC view.

Under primitivism, emotion qualifies as an investigative kind because every instance is explained by a common underlying mechanism: the interaction between the general affective system and exteroceptors. At the same time, emotion qualifies as a natural kind because every emotion is defined according to a specific essence: emotions are affective states that carry information about the organism–environment relationship. However, things are different when it comes to classifying discrete emotions as natural and investigative kinds. Some emotions may be investigative kinds, because evidence suggests that at least some

emotional experiences are triggered by reliably reoccurring emotional reactions (i.e., patterns of bodily change). But that does not license a claim that those emotions also are natural kinds, since (per essentialism) there is no principled means of distinguishing one emotion's affective quality as unique to its category. And if no unique essence can be found, then that category does not qualify as a natural kind.

Chapter 1

What is a Jamesian emotion?

Introduction

It is widely acknowledged that William James helped to shape the course of psychology as an autonomous discipline, one distinct from both biology and metaphysics. Moreover, his theory of emotions is also recognized as historically significant. Still, his claim that emotions consist in feelings of bodily changes (James 1884; 1890, II: ch. 25; 1992, ch. 24; 1894) is typically discarded by contemporary emotion researchers. This chapter seeks to reverse that trend and to make a case for rehabilitating James on emotion.

As regards his critics, the physiologically inclined disagree that the parts of the body James identified as contributing sensations to emotional experiences can operate in the way he claims. And yet, they seem unable to agree on exactly which parts of the body James was talking about in different iterations of his theory. They also seem to miss James's distinction between mere "stirrings of the body" and an organism's behaviors. These contribute differently to our emotional experiences, yet this point of James's is paid little attention today.

Psychological obstacles also stand in the way of his rehabilitation. Appraisal theorists complain that his theory is anticognitivist when in fact emotions clearly involve cognitions:

Magda Arnold, the first theorist to explicitly posit an appraisal mechanism as responsible for emotional experience, states that “James faced the problem that some such recognition or appraisal is necessary between the cold perception [of a stimulus] and the physical changes [of an emotion], but did not solve it” (Arnold, 1960a, 107; cf. Averill, 1992, 223). Such criticisms take aim at James’s story of how emotions are triggered and experienced.

The problems don’t end there. Philosophical issues surround the intentionality of emotions on any given picture; on James’s theory, his critics claim, emotions lack intentional content and are, therefore, rendered functionally useless to the organism (“merely epiphenomenal”), contrary to our intuitive acquaintance with them in everyday life. (It’s not just that we feel fear when some danger looms, they say; it’s that we’re scared *of* that looming danger.) Contemporary emotion theorists gravitate towards accounts that posit cognitive causes (such as appraisal mechanisms) for emotional experiences, since those causes provide a means of securing emotions’ intentional aspects. James thus comes under attack from all sorts of angles. It’s hardly surprising, then, that his theory of emotions has faced a rather cold reception since its initial publication more than a century ago.

Although the centenary of his magnum opus, the *Principles of Psychology*, occasioned a number of celebratory papers revisiting his theory,³ their collective impact has not fully reinvigorated theoretical interest in Jamesian emotions. I think that James is treated unfairly these days. This conviction motivates this chapter, which is to examine whether commentators actually analyze James’s theory of emotion when offering assessments of it. I will argue that commentators typically misunderstand James’s theory because they fail to analyze adequately the concepts from which it is built. In particular, they tend to interpret James’s terms *perception* and *cognition* as they are described in current parlance. This leads commentators to put forward, often prematurely, one of three general conclusions regarding the Jamesian

³See for instance Margaret Donnelly’s (1992) edited collection *Reinterpreting the Legacy of William James*, and the papers comprising volume 1, issue 3 of *Psychological Science* (1990). Several of these commentaries are discussed below.

theory of emotion:

- (C1) it is incompatible with our current (mainstream, presumably more accurate) understanding of the mind and therefore false;
- (C2) it is compatible with our current understanding of the mind and we just need to tease out of it the implied elements that render it so; or
- (C3) it can be made compatible if we tweak it a bit and supplement it with psychological mechanisms identified in our current scientific theories.

To follow the first path is rather imperious: it classifies James's theory as unacceptable from the standpoint of contemporary conceptual demands, without the benefit of a detailed comparative assessment. The second route brings with it the danger of reading the present into the past in the process of "teasing out" what is valuable in James, thereby misrepresenting and misapprehending his place in the history of the discipline of psychology. And the third move overly treats present theories as saviors of the past; it assumes that the history of psychology simply progresses linearly, towards accuracy of explanation, with a general tendency for concepts and conceptual frameworks to endure (with permutations) rather than being replaced or challenged by alternatives. But anyone who buys that a "cognitive revolution" dragged psychology out of the Dark Age of Behaviorism should be disinclined to espouse such approaches without having very carefully examined the overlap of ontologies between contemporary theories and their historical counterparts. Cognitive psychology is the product of a conceptual revolution, just as is behaviorism, and James's psychology was, too. One cannot neatly track concepts across these three distinct ontological periods without first doing some solid conceptual and contextual analysis.⁴

⁴James's theory of psychology generally relies on introspection to ground descriptions of mental and other cognitive events. Behaviorism, in contrast, favors explanations of behavior based on observable correlations with varied extraorganismic stimuli. For behaviorists such as James Watson (1930) and B. F. Skinner (1953), the inner workings of the human mind are at best secondary to stimulus-response pairings in psychological explanations; just as often they are shunned from the behaviorists' vocabulary. Early cognitive psychologists

Of course, it may turn out that one of these three conclusions is correct. The aim of this chapter is not to prove that James's theory of emotion is right. It would be utopian to assume that current psychological science is in a position to rule on such a matter. Rather, this chapter aims to gather maximal evidence of internal consistency. Historians of philosophy and psychology tend to promote charitable, complete understandings of important works prior to deeming them either plausible, implausible, or necessarily incorrect. And if it turns out that James's theory of emotion, properly understood, is plausible, then it might guide empirical investigation. For these reasons, James deserves a fair trial. Therefore, the question to be answered is this: once we situate it within his broader theory of psychology, what does James's theory of emotions claim?

This chapter, in elucidating the theory, offers cases in which commentators misconstrue it and so prematurely conclude of James's theory either (C1) or (C2). Since (C3) involves modifying rather than clarifying James's theory, I delay considering it until Chapter 2. Section 1.1 gives an overview of James's theory, making note of his ambitions and the larger project from which the theory is developed. It then highlights two common ways in which James's theory is misconceived: by treating it as identical with the so-called James-Lange theory, which in fact has never existed as a unified theoretical proposal; and by failing to recognize the way in which the bodily changes productive of emotional experiences relate to James's account of instincts. Both of these misconceptions invite us to conclude that James's theory is incorrect: the first strips Jamesian psychology of its personalist foundation; while the second ignores a crucial distinction between bodily changes and world-directed behaviors. When delivered from these two misconceptions, James's theory is secured against those who would quickly dismiss it on conceptual grounds. Section 1.2 discusses an interpretation of James that requires a cognitive appraisal mechanism to be at work in triggering the bodily changes responsible for our emotional experiences. This interpretation is incorrect because

revived explanatory interest in the internal machinations of organisms' cognitive systems and incorporated concepts of mental representations into causal descriptions of psychological phenomena.

it requires accounts of perception and cognition that are contrary to James's own. Instead, Jamesian emotions can satisfy common demands made of a theory of emotion without having to resort to the sorts of cognitive machinery typical of contemporary theories. I conclude the chapter by pointing out some practical benefits of a faithful reading for present-day theorizing.

1.1 Standard interpretations and misinterpretations of James

This section presents what may be called the “standard interpretation of James”. It begins by outlining the key claims and hypotheses of James's theory of emotions (section 1.1.1). Section 1.1.2 discusses two ways in which James's theory is often misconstrued, and why these misconstruals are harmful to understanding his project.

1.1.1 James's theory of emotion: key claims and hypotheses

While the most widely cited iteration of James's theory of emotions appears in a stand-alone article titled “What is an Emotion?” published in 1884, this article was composed as part of a much larger and more complete project: the two-volume opus the *Principles of Psychology*, eventually published in 1890 (James, 1890, I & II; hereafter *Principles*). James developed this project to account for psychology as a discipline: to map out the classes of psychological phenomena in terms of their relationships to one another and to other domains, including brain science. He often outlines his formal understanding of key concepts before employing those concepts in his treatment of subsequent topics. The fundamental view I adopt in this chapter is that James's theory of emotions is most clearly understood when situated within the larger conceptual framework in which it was composed. His theory of emotions cannot neatly sit alongside a contemporary ontology informed by the information processing

models of mind proposed or motivated by traditional cognitive psychology of more recent times. Conversely, it can succeed quite well within its proper environment.

The version of the theory typically cited in literature on emotions takes a very simple form. James argues that emotions are feelings arising when the mind perceives changes across the body due to the body's being excited by some perceived stimulus (James, 1884; 1890, II: ch. 25; 1992, ch. 24). This claim reverses what he calls the "common sense" view of the emotion sequence, which holds that emotional feelings lead to behavioral changes: "we lose our fortune, are sorry and weep; we meet a bear, are frightened and run; we are insulted by a rival, are angry and strike" (James, 1884, 190; 1890, II: 449-450; 1992, 352). Instead, James writes, it's the physiological changes that bring about the feelings of emotion:

The hypothesis here to be defended says that this [common sense] order of sequence is incorrect, that the one mental state is not immediately induced by the other, that the bodily manifestations must first be interposed between, and that the more rational statement is that 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, as the case may be. (James 1884, 190; 1890, II: 450; 1992, 352)

We can grossly characterize his preferred causal schema like this:

STIMULUS → PERCEPTION OF STIMULUS → BODILY EXCITEMENT → EMOTIONAL FEELING

At least, this characterization covers what James calls the "standard" (1884) or the "coarser" (*Principles* onwards) emotions: "cases in which a wave of bodily disturbance of some kind accompanies the perception of interesting sights or sounds, or the passage of the exciting train of ideas" (James, 1884, 189). The three cases mentioned above are representative of these coarser emotions. James considers coarser emotions to be paradigmatic of the whole class and most in need of explanation. The above characterization captures only the broadest

details of James's theory. Each element therein requires careful analysis and enrichment so that James's overall position is accurately understood. The remainder of this chapter offers exactly these analyses.

Note that what James proposes is a theory according to which an emotion is a conscious feeling. Emotions do not *include* bodily changes; they are caused by those changes. The excited body no more constitutes the emotion itself than does the stimulus that triggered that excitement—be it a financial loss, or a bear, or an insult, or whatever. The most important, immediate cause of an emotional feeling is a set of bodily changes, and “*our feeling of the same changes as they occur is the emotion*” (James, 1884, 190-191, original emphasis).

1.1.2 Misinterpretations and confusions

James often uses the term *emotional reaction* to refer to the set of bodily changes that produce an *emotional experience*, the feeling of an emotion. Moving forward, this chapter employs these two terms discretely to refer to their respective phenomena: emotional reactions are physiological, while emotional experiences are subjective feelings. Although James considers only emotional experiences to be emotions proper, he is aware that a complete scientific picture of emotions must take into account their antecedent causes. Emotional reactions earn their label by virtue of their direct relationship with emotional experiences. Nonetheless, it is important to keep these two concepts distinct from one another. The next subsection offers two cases of common misinterpretations that ignore this difference, to their detriment. Once we clear away these misconceptions, James's theory becomes a more plausible guide to empirical research.

1.1.2.1 James but not James-Lange

In the year after James's article appeared, Danish physiologist Carl Lange published *Om Sindsbevaegelser* (Lange, 1885), a psychophysiological study of the emotions. James notes

significant similarities between his own and Lange's theories. He quotes the latter at length in his chapter in the *Principles*, citing Lange's physiological claims as evidence in his favor. And in his 1894 article "The Physical Basis of Emotion" he says that he and Lange developed "the same theory of emotional consciousness" (James, 1894, 516). Ever since John Dewey formally conflated the theories in 1894, and especially since Walter Cannon's forceful criticism of the conflated view in 1927, James has become known as one half of the contributing force behind the "James-Lange theory".

Here I am concerned with James's theory but not the James-Lange theory. The reason is simple: despite James's academic grace in acknowledging a like-minded theorist, there are significant differences between his and Lange's theories of emotion.

One point of departure is in the physiological details of either theory. James and Lange hold distinct beliefs regarding which parts of the body are involved in emotional reactions, the changes that cause emotional experiences in consciousness (see also Izard, 1977, 55; Mandler, 1984, 20; and Prinz, 2003, 70). Lange's descriptions of emotions involve clear physiological details—far more so than do James's. He writes: "We owe all the emotional side of our mental life, our joys and sorrows, our happy and unhappy hours, to our vasomotor system" (Lange, 1922, 80), whose disturbances serve as "the only primary symptoms" of emotions (Lange, 1922, 60). This tight circumscription of physiological factors at play in emotions is central to Lange's theory of emotion, but, as we will see below, it did not sit well with James.

Lange also offers physiological and physiognomic descriptions of specific emotions. Like James, he is amenable to the idea of discounting common-sense emotion concepts: physiological observation should inform our understanding of specific emotions as much as it guides our understanding of emotionality more generally. Accordingly, he describes the physiological events that best represent numerous emotions we readily identify in life. Sorrow is readily observable in its "inhibitory effect upon the motor apparatus" and hypoemia throughout the body (Lange, 1922, 40-41), with a lack of blood supply to the brain account-

ing for “mental lassitude, dullness, a feeling of mental fatigue and effort” (Lange, 1922, 43). Fear or “fright” is “closely related to sorrow”, sharing most of the physiological characteristics but manifesting them in “a more intense or acute form”, with some added convulsions of the voluntary muscles (Lange, 1922, 48). Anger and joy, one would think, are easily distinguishable from one another—at least from the first-person perspective. Lange, however, notes that observing physiological manifestations in either case offers far fewer means of distinguishing between the two than one might expect: save for some widening of the blood vessels in the face during anger, distinguishing symptoms are opaque (Lange, 1922, 49-50). Anger and joy, in contrast with sorrow and fear, see an increase of blood to the brain where it facilitates speedy mental functioning (Lange, 1922, 45). For Lange, then, emotions are physiologically discrete events.

This contrasts with James, who emphasizes the variability and dynamism of emotional reactions (see Gendron & Barrett, 2009, 323-325). James is critical of Lange’s claim that anemia in the brain is principally responsible for feelings of grief (James, 1890, II: 446), and of his rigid explanation of grief’s other constituent bodily changes (James, 1890, II: 444). He further criticizes Lange’s overreliance on “the vaso-motor factor in his explanations” while defending himself against Wilhelm Wundt’s attack on the James-Lange theory (James, 1894, 517).

Though James clearly disagrees with Lange about the physiology underlying emotional experiences, there is little consensus in the secondary literature on which parts of the body he identified as key contributors to emotional experiences. Richard Davidson (1994, 237) and Jesse Prinz (2003, 70) state that Jamesian emotions involve skeletal muscles; Antonio Damasio (1999, 288) says James gave skeletal muscles “short shrift”; while Phoebe Ellsworth (1994, 225) accuses James of wavering on the issue. Most commentators point out James’s emphasis on visceral organs as sources of felt bodily changes, but they interpret differently the extent to which he emphasizes the viscera (see, e.g., Ellsworth, 1994; Izard, 1977, 55-56;

and Lang, 1994). Other commentators (e.g., Freeman, 2000, 215; Griffiths, 1997, 80) offer a looser, more contemporary description of James's identified bodily changes, couching them in the autonomic nervous system—which is defined by bodily functions, not by anatomical regions.⁵ And we've just seen that James includes only limited vasomotor activity. I will not attempt to clarify James's physiological claims in this chapter. As received interpretation goes, this lack of physiological clarity on James's part makes mediating between James and Lange more difficult, but it also renders illegitimate an assumptive conflation of the two theories.

From a more abstract perspective, however, one could argue that their physiological disputes are minor empirical disagreements, and that James and Lange still agree on the grander theoretical claim regarding the relationship between bodily changes in general and emotional experiences. Nonetheless, Peter Lang argues that Lange's position on emotions “seems to have a separate emphasis” from James's: while James is interested in explaining emotional consciousness (remember, emotions are feelings), Lange takes emotions to be constituted by physiology; conscious experiences thereof are of secondary interest (Lang, 1994, 212). After detailing the physiological characteristics of the particular emotions mentioned above, Lange asks:

Is it possible that vasomotor disturbances, varied dilation of the blood vessels, and consequent excess of blood, in the separate organs, are the real, primary effects of the affections, whereas the other phenomena,—motor abnormalities, sensation paralysis, *subjective sensations*, disturbances of secretion and intelligence—are only secondary disturbances, which have their cause in anomalies of vascular innervation? (Lange, 1922, 58, emphasis added)

His answer, stated a few pages later, is affirmative: “the emotion consists exclusively of these

⁵The autonomic nervous system was named by J. N. Langley in 1903, so James didn't have access to this term in his writings on emotions.

functional disturbances of the body” (Lange, 1922, 64). For Lange, emotions are physiological, not psychological, phenomena. Emotional experiences take the back seat in his explanatory project as downstream effects of emotions proper.

Distinguishing the two theorists’ projects brings forward the implications of taking physiological or psychological phenomena as central to one’s theory. In James’s case, understanding emotions as feelings invites scientific inquiry into subjective experience. Jamesian psychology is fundamentally personalist, meaning that a notion of *self* is central to an understanding of the mental. In fact, Hazel Markus has flagged the self as “perhaps *the* basic problem” of interest to James (Markus, 1990, 181, original emphasis). His hundred-page chapter on consciousness of the self, in the *Principles*, is by far the longest across the two volumes. Therein he describes the self as *felt*. He offers an introspective report of his own feeling of self, describing it in much the same way as he does emotional experiences:

In the first place, the acts of attending, assenting, negating, making an effort, are felt as movements of something in the head. In many cases it is possible to describe these movements quite exactly. In attending to either an idea or a sensation belonging to a particular sense-sphere, the movement is the adjustment of the sense-organ, felt as it occurs. I cannot think in visual terms, for example, without feeling a fluctuating play of pressures, convergences, divergences, and accommodations in my eyeballs. (James, 1890, I: 300)

After describing the phenomenality of remembering, reflecting, consenting, negating, and exerting effort, he draws his conclusion about the physiological cause of these self-experiences: “In a sense, then, it may be truly said that...*the ‘Self of selves,’ when carefully examined, is found to consist mainly of the collection of these peculiar motions in the head or between the head and throat*” (James, 1890, I: 301, original emphasis). This “Self of selves” is the phenomenological experience of one’s self as a probing, attentive agent. And although these feelings might be quite easily located or identified in the body, James also speculates

that the experience of one's "spiritual" or "inner" subjective self "is really a feeling of bodily activities whose exact nature is by most men overlooked" (James, 1890, I: 301-302, emphasis removed).

Conscious experiences are thus intrinsically related both to the self and to the body. A theorist cannot explain them in purely objective terms—at least, not nearly as neatly as she can physiological changes. James is no eliminativist about psychology; in fact he even argues that his theory of emotions should not be considered "materialistic" (James, 1890, II: 453; 1992, 356-357). Yet the personalist feature of Jamesian psychology seems to be lost in contemporary readings of James's emotion theory. James Averill offers an autopsy of James's "dead end" theory of emotions that in fact reverses James's considered view of the relation between emotion and felt experience: "...by focussing attention on bodily changes during emotion, [James's theory] has led investigators to ignore... [that] emotions are intimately related to a person's sense of self, both in terms of eliciting conditions and consequences..." (Averill, 1992, 222). This claim is directly contradicted by James's own writings on the subject. In James's theory, one's emotional experiences and one's feeling of self both flow from the same spring.

If the personalist side of Jamesian psychology is too often overlooked when reading James alone, then confounding his theory of emotion with Lange's theory scuttles the self entirely. Neither an explanatory framework nor a particularly fruitful guide to empirical investigation is likely to derive from our treating James's and Lange's theories as unified in any rich sense.

Instead, when James's and Lange's theories are erroneously equated a strawperson becomes available for critics. Kurt Koffka, for instance, notes that "the only real achievement of the [James-Lange] theory is its insistence upon the fact that emotions are more than just contents of consciousness" (Koffka, 1935, 401). Worse still, William Lyons writes James's legacy in a behaviorist's ink when he says that John Watson's account of emotion "is really

William James's account shorn of any reference to feelings or any other sort of conscious state" (Lyons, 1999, 32).⁶ Explicitly or implicitly, both of these comments take James's definition of emotion to include bodily changes. This misrepresents James thoroughly. His theory holds that if you take away the conscious feeling you eliminate the emotion entirely. One cannot use James's theory to explain emotional reactions, since the constitutive bodily changes are antecedent to the emotions proper (see also Mandler, 1984, 18). Even if Koffka and Lyons were to defend themselves in saying that they employ the term *emotion* according to their own definitions (and not James's), in attributing these views to James they strip Jamesian emotions of their personalist bases. In conceptual cases such as these, conflation leads to confusion: "Thus, the James-Lange theory is cited variously both in support of and against physiological reductionism, as the precursor of the behavioral analysis of emotion, and as the starting point for cognitive theories of affect" (Lang, 1994, 212).

Collectively, these disagreements over the physiological minutiae underlying emotions, along with how one identifies a phenomenon of interest, weigh heavily in favor of disbanding the James-Lange theory and focusing on each theorist individually. At the most general level of analysis the distinctions may seem superficial, and this is probably the level at which James himself was equating the two theories in his 1894 discussion. But for the sake of both historical veridicality and progression of an empirical research program—specifically, in guidance of experimental investigations to test hypotheses—such details are crucial.

1.1.2.2 "Afraid because we run": emotional and instinctual reactions

Having disentangled James's theory from the James-Lange postulate, we are now in a better position to examine the nature of *emotional reactions*, the physiological changes triggered by stimuli. Given the lack of consensus regarding James's position on the physiology of emotional reactions, it is not surprising that there is confusion regarding how to read James on

⁶Watson himself made no such mistake in interpreting the spirit of James's theory before scathingly attacking its introspectionist foundations (see Watson, 1930, 141-142).

this topic. Nonetheless, James does provide a principled means of distinguishing emotional reactions from other kinds of actions according to their functional aspects and their downstream psychological effects.

Unfortunately, this distinction is consistently overlooked by James's commentators. They often paraphrase incorrectly James's positive claim that we are "afraid because we tremble" as the very different claim that we are "afraid because we run"—the latter a claim James never endorsed. This was a problem James noticed in his own time, and, despite the burden being placed on the shoulders of the careful reader, James took responsibility for it wholeheartedly: "I admit that my own text set a bad example when it said 'we are frightened because we run'" (James, 1894, 519). In fact, this is the solitary instance, throughout his corpus, in which he writes the problematic phrase. He then clarifies what he meant: we should "let the word 'run' but stand for what it was meant to stand for", namely certain patterns of changes across the body, many of which may be "invisible" and visceral, that are felt as the emotion (James, 1894, 519). It's interesting that James would misquote himself (by suggesting he had used the phrase elsewhere), but it demonstrates that his original ambition was to overturn thoroughly the "common sense" theory prevalent in his era. Forceful, somewhat hyperbolic language was his means of emphasizing the novelty and radicalness of his theory. Such sentiments are echoed in a retrospective paper by Edmund Delabarre, one of James's students (and a contributor to the *Principles*: see James, 1890, II: 13-27), who writes:

We appreciated fully his [James's] remarkable genius for felicitous, clear and picturesque expression; although occasionally this led to complete misunderstanding of his meaning,—as when he said, in expounding his famous theory of emotions, that "we are sorry because we cry, afraid because we run," not the other way round. Evidently, we can be sorry without crying, afraid without running. (Delabarre, 1943, 125-126)

James's self-defense often goes ignored—further evidence that commentators fail to en-

gage properly with his corpus. He is still read as saying that we're afraid because we run. Within the contemporary literature, guilty parties include Robert Plutchik (1970, 3), James Laird and Charles Bresler (1990, 641), Robert Zajonc (Zajonc et al., 1993, 209), Tim Dagleish (2004, 582), Asher Koriat (Koriat, 2006; Koriat et al., 2006, 37), Bruce Friedman (2010, 384), and Gerald Clore and Andrew Ortony (2013, 338). Joseph LeDoux is a repeat offender (LeDoux, 1996, 84-85; 2002b, 201; 2007, 395-396; 2008, 86; Fellous et al., 2003, 398; Fellous & LeDoux, 2005, 80). This false attribution can now be found throughout a number of psychology and psychiatry textbooks, as well as in online encyclopedia entries on James.

Getting James right on this point is important. Running and trembling are very different actions with distinct functions: they represent different ways of interacting with the world. Thus, their relations to emotions dictate what kinds of functional roles emotions play in Jamesian psychology. To class the difference in Jamesian terminology, running is an *instinctual reaction* while trembling is an *emotional reaction*. The latter “fall short of instincts, in that the emotional reaction usually terminates *in the subject's own body*, whilst the instinctual reaction is apt to go further and enter into practical relations with the exciting object” (James, 1890, II: 442, emphasis added). The emotional reaction is so named because it is that subset of instinctual bodily changes that cause a “diffusive wave” of nervous activity, thereby giving rise to the characteristic feeling of an emotion. Thus, the trembling itself is not fear proper, but its cause.

Yet the trembling can serve as a catalyst for a world-directed response such as fleeing from or fighting off the perceived threat. Following Herbert Spencer's (1855) and Charles Darwin's (1872) leads, and in line with many ecologically minded, evolutionary theorists of emotion today (e.g., Ekman 1992b, 1999; LeDoux 1996; Griffiths 1997), James sees the bodily stirrings of emotional reactions as reflexes typically operating in service of the well-being of the organism (James, 1890, II: 477-483; 1992, 362-365). Recent empirical evidence favors this ecological view of emotional function: for instance, changes to the facial expressions

during a fear reaction cause an enhancement of the upper visual field and increased nasal respiratory capacity, while the feature changes of disgust show the opposite ecological effects (Susskind et al., 2008). As I'll explain shortly, the best interpretation of the relationship between emotional reactions and instinctual reactions is that the latter are often triggered by the former. If, however, we cannot find a nice teleological story of why certain bodily changes occur during instinctual responses to stimuli, nothing is lost by acknowledging that such responses occur. Those “purely idiopathic effects of the stimulus” that cannot be explained by appeal to teleofunctional principles can instead be described as “purely mechanical or physiological” permutations (James, 1890, II: 482).⁷

Moreover, overt behavior such as running is not a necessary effect of emotional reactions, and emotional reactions can be entirely unobservable (hence “invisible”) to anyone but the emoter: consider how well a deer can hide its fear by staying perfectly still upon encountering a predator. Another way to put James's point is that emotional reactions can extend into instinctual behaviors. In terms of causal history, their distinction is not of kind, but of degree. James tells us as much when he states that it is “quite arbitrary from the scientific point of view” whether the physiology of fear be described as an instinct or an emotional reaction (James, 1992, 350); elsewhere, he writes that emotional reactions and instinctual reactions “shade imperceptibly into each other” (James, 1890, II: 442).

James also offers reason to distinguish between emotional and instinctual reactions on phenomenological grounds. To return briefly to his position on the feeling of self, James asks us to consider an implication that would follow from his theory, as outlined above in section 1.1.2.1, if it were true. We can make a conceptual division between two kinds of physiological acts: *adjustments* and *executions*. Feelings of adjustments ground a felt “nuclear self”, while feelings of executions—actions performed by the body toward the world—contribute to the phenomenality of “the less intimate, more shifting self, so far as it [is] active” (James,

⁷Darwin offers a similar principle in his *Expressions of the Emotions*, according to which excessive nerve-force is generated when the sensorium is strongly excited (Darwin, 1872, 35).

1890, I: 302). This distinction between physiological adjustments and executions looks to be identical to his distinction between emotional and instinctual reactions. If so, then both emotional experiences and the enduring, intimate experience of one's self are caused by emotional reactions. Conversely, performances of instinctual reactions contribute less phenomenal character to one's sense of self. Therefore, emotional and instinctual reactions are differentiated by both their functions and their psychological consequences.

In the sense that he failed, in his 1884 article, to draw the distinction between emotional reactions limited to bodily stirrings (*adjustments*) and world-directed instinctual reactions (*executions*), James is to blame for bringing this general interpretative problem upon himself. Though he never claimed that we are frightened because we run, he did say that we are angry because we strike. Here, he should have said that we are angry because we feel the changes our bodies make in preparing to lash out. With that said, one cannot rule out that running due to a fear reaction would add to or reinforce that fear reaction, and subsequently maintain and amplify one's felt emotion of fear. Nor can we rule out that actually striking out would produce or strengthen a feeling of anger (cf. Laird & Bresler, 1990, 641-642). The important point here, again, is that overt behavior is unnecessary for emotional experience. Emotional reactions and downstream behavioral responses are intimately connected, but it is only the stirrings of the body that are the direct, necessary, and sufficient triggers of emotions proper.

This section has described two misconceptions of James's theory that have damaged its perceived plausibility. When read into his theory, they invite its swift dismissal. That is, these misconceptions provoke the conclusion labeled (C1) in the introduction: James's theory of emotions is incompatible with our current understanding of the mind, and is therefore false. I have argued that James's claims are quite different than these readings would lead us to believe. Having rebutted them, we cannot be so quick to reject his theory.

1.2 Are Jamesian emotions epiphenomenal?

Probably the most common dismissal of James's theory of emotions comes from philosophers and psychologists who take emotions to be cognitively richer phenomena than James has proposed. Emotions are described variously as intentional, evaluative, perhaps even propositionally constructed cognitive states whose causes and experiential characteristics cannot be exhausted by physiology or sensations. In reducing emotions to mere sensations arising as the outputs of a process involving perception and reflex responses, it seems that such qualities, which other theories build into emotional phenomena, are necessarily missing from James's picture (Damasio, 1994, 129-130). That is, it looks as though James renders emotions epiphenomenal (Deigh, 2001, 1252; 2010, 20-21). If they are understood as such, John Deigh tells us, "the importance of emotions as an object of study diminishes considerably" (2001, 1252; see also Solomon, 2003, 12). And it's easy to understand why. According to both ordinary intuition and the vast majority of theories past and present, emotions must be understood as *information-bearing*, *intentional*, and *motivational* phenomena.⁸ If Jamesian emotional feelings are epiphenomenal then they fail to provide any of these three crucial features. Little wonder, then, that a contemporary commentator like Robert Solomon would think it is "pathetic" that the likes of James have convinced members of the field to embrace non-cognitivist positions (Solomon, 1980, 271).⁹

Cognitivist theories of emotion are engineered to incorporate into their explicanda those psychologically impactful functions mentioned above. Recall that cognitivists understand emotions either as judgments or responses to judgments. Versions of cognitivism far pre-date James (e.g., Aristotle and the Stoics) and have not much loosened their grips on the

⁸This set of three desiderata for a theory of emotion is an identifiable trend throughout the history of the literature, at least in Western scholarship. Barring the stripped-down conception championed by behaviorism, emotions are always conceived of as responses *to* something—a feature of the world or of the self—and as such they involve an intentional aspect and a motivational aspect. The idea that emotions are information-bearing is found at least as early as the Stoics, who conceived of them as judgments.

⁹By the way, the word pathetic originally means "arousing the emotions".

field since—especially given the birth of cognitive psychology in the mid-twentieth century. An extreme iteration of cognitivism is offered by Martha Nussbaum (2001; 2004), who argues that emotions are judgments about our well-being (hence cognitions) that need not include affective qualities such as the emotional experiences James describes. Although we humans *do* have such feelings when we have emotions, this fact is merely a contingency due to our being embodied minds. On the other hand, “a thinking being...could care deeply about something in the world, and have the thoughts and intentions associated with such attachments. And that’s all we really require for emotion” (Nussbaum, 2001, 60; cf. 2004, 28).

In this section I explore one way in which emotion theorists have attempted to construe James in light of some of the demands mentioned above. According to this *appraisalist interpretation*, James was a cognitivist all along—he just didn’t make that clear given his old-fashioned use of the term *cognition*. This reading generates a version of (C2) as characterized in the introduction: the idea that we can make James’s theory compatible with contemporary views by reconstructing it using contemporary terminology. The appraisalist interpretation is outlined in section 1.2.1. Section 1.2.2 provides evidence that it is interpretively incorrect, and section 1.2.3 shows that James’s theory need not employ an appraisal mechanism in order to possess those three features demanded of emotions: the bearing of information, intentionality, and motivation.

1.2.1 An appraisalist interpretation

Some interpreters attempt to clarify James’s position by elaborating on aspects they believe the author himself neglects. Phoebe Ellsworth (1994), for instance, dresses James’s theory up in more contemporary terms, teasing out of it a process of appraisal at work in the early stages of an emotion sequence. Laird and Bresler (1990, 646), Richard Lazarus (1991, 39-40), E. Virginia Demos (1992, 212), Klaus Scherer (1996; 1999; 2001), and Jenefer Robinson

(2010; cf. 2005, 28-29) also read James as covertly positing an appraisal mechanism. Appraisals, according to appraisal theory generally, are evaluative judgments of one's situation that provide information about a potential or actual alteration of one's well-being.¹⁰

There is much debate over an appraisal mechanism's status as a cognitive system. Numerous proposals for how such judgments are formed describe cognitively simple mechanisms, such as Arnold's instinctive "sense judgment" view (Arnold, 1960a, 170-177). These can be compared with Lazarus's more sophisticated dimensional appraisal system, whereby perceptual content is assessed according to the criteria of "*goal relevance, goal congruence or incongruence, and type of ego-involvement*" (Lazarus, 1991, 133, original emphasis). All such views take appraisal processes to evaluate stimuli in relation to the self. Lazarus illustrates this point:

What would transform sensory states into emotions? The transformation necessary to produce an emotion out of sensory states is an appraisal that those states are favorable or damaging to one's well-being. When we cognize an event as pleasant or unpleasant, we are not experiencing an emotion. However, when we further cognize that we are or may be personally benefited or harmed, the cognitive transformation has gone beyond the mere registration of discomfort, and the experience becomes an emotion. (Lazarus, 1984, 126)

For the most part, appraisal theorists are cognitivists about some aspect of an emotional episode, though they might disagree over where meaning enters the picture: psychologists often propose cognitive causes for emotions, while philosophers take cognitivism about emotions to imply that emotions themselves are cognitions—that is, they are intentional, information-bearing mental states directed at the environment (Reisenzein & Döring, 2009, 199). Regardless of one's conception of where *cognition* enters in, appraisalists agree on the

¹⁰For a good introduction to the general claims made by appraisal theorists, see Ellsworth & Scherer (2003); for a brief history of appraisal theories, see Schorr (2001).

idea that sensory information is processed in some way, inside the organism, to produce a value judgment about what is sensed with respect to that organism. Fear, on this picture, might result from a looming object's being appraised as dangerous. This idea imbues an episode of fear with meaning.

Ellsworth and others offer an *appraisalist interpretation* of James's standard theory. Her aim is to show that James did not leave various gaps in his explanation, as Arnold would have it (as quoted on page 19). If James is an appraisal theorist, then, the appraisal of danger is what triggers a set of bodily changes that provides the sensation of fear.

Ellsworth reads James as implying that the kind of stimulus perception usually involved in triggering emotions is just a form of cognitive appraisal; she thinks that there is "no question" that this is the case (Ellsworth, 1994, 224). Lazarus thinks the same, stating that James made a mistake in using the word *perception* instead of *appraisal* (Lazarus, 2001, 39). Laird and Bresler even say that James took one of the merits of his theory to be its shifting attention to the question of why "particular appraisals might induce" particular kinds of emotional reactions (Laird & Bresler, 1990, 646). To Ellsworth, James's thought that cognitive appraisal is at work in triggering emotions "is so obvious that it needs no special emphasis" (Ellsworth, 1994, 223); instead of mentioning it, James focused his reader's attention on "what he thought was his most original idea—the role of bodily processes" (Ellsworth, 1994, 224).

Note that this rendering significantly complicates the simple picture of James seen throughout the literature on emotions (and characterized on page 23). The immediate causes of an emotional experience are still bodily changes, so Jamesian emotions (*qua* conscious feelings) remain noncognitive phenomena. But what causes those changes is explained according to a mediative, *cognitive* perceptual mechanism and not a reflexive, automatic response system. This gives us the following appraisalist construal of James's theory:

STIMULUS → COGNITIVE APPRAISAL OF STIMULUS → BODILY EXCITEMENT →
EMOTIONAL FEELING

Ellsworth actually goes further than this and claims that the feeling of bodily changes combines with the perception (appraisal) of the object to form an emotion (Ellsworth, 1994, 223). This further complicates the sequence, and indeed makes a Jamesian emotion more than mere feeling of bodily changes (but still, *contra* Koffka's reading, not more than the contents of consciousness), since the emotion is now simultaneously a perception of the stimulus object and a perception of one's bodily reaction. In this way, emotions come to possess a world-directed intentional aspect, thereby satisfying one of the intuitive desiderata. But this interpretation runs counter to James's explicit statement, repeated in triplicate (James, 1884, 189-190; 1890, II: 449; 1992, 352; quoted on page 24) and reworded in the opening paragraph of his 1894 discussion article: emotions *just are* feelings of bodily changes as they occur. Ellsworth's rather radical interpretation of James, which posits cognitive appraisals as both initiators *and* constituents of emotions, has been convincingly refuted (see Reisenzein et al., 1995). I will therefore refocus on Ellsworth's more restrained (yet still contentious) interpretation that sees appraisals as catalysts for the bodily changes that produce emotional feelings.

Before assessing Ellsworth's reading, we should consider what it implies should it be correct. What work does the purported appraisal mechanism do for James's theory? At a first pass, it appeals because it helps to safeguard those features of emotion that intuition demands. Firstly, the appraisal mechanism provides information about the current state of affairs with respect to the organism's ecological concerns. An appraisal of a looming, growling bear-shape will inform the organism of a present danger to which it must react by, say, running away. In providing information about a danger, the appraisal generates an evaluative judgment. Secondly, since this judgment is about the present scenario, it possesses the requisite feature of environmentally directed intentionality. Finally, the intentional information is likely responsible for triggering the bodily changes that motivate the behavior of running away. In short, it looks as though building an appraisal mechanism into the Jame-

sian emotional sequence would see emotions satisfy the three intuitive desiderata.

However, the appraisalist interpreter should not celebrate too quickly. In fact all of these feature manifestations precede the emotion itself. Remember James's words: the emotion *just is* the bodily feeling. Deigh's charge of epiphenomenalism stands, since the appraisalist James still doesn't provide us with a function for emotional experiences. Of course, this issue is really about what counts as "an emotion", and whether the episodic events causally preceding an emotional experience count as constituents of the emotion itself. We're faced with two choices if we accept Ellsworth's interpretation: either we accept that emotions (*qua* conscious feelings) are potentially epiphenomenal and therefore psychologically uninteresting, or we diverge from James's use of the term *emotion*, widening it to apply to the whole episodic sequence formulated on page 38.

Note that the first choice only forces us to accept that emotions *may* be epiphenomenal. We would need an independent argument, from James or made on his behalf, that emotional experiences perform some function.¹¹ Yet whether James's theory is correct should not bear on whether one's interpretation of James is correct. Therefore, we should assess Ellsworth's interpretation according to the degree of accuracy with which it rebuilds James's theory. Section 1.2.2 argues that Ellsworth's reading should be rejected. Section 1.2.3 offers some reasons for thinking that James's theory can satisfy the intuitive desiderata.

¹¹Such an argument would come less easily for James than for proponents of the common sense view he targets. Darwin, for instance, assumes that "most of the expressions and gestures involuntarily used by man and the lower animals"—that is, what I've been calling emotional reactions—are brought about "under the influence of various emotions and sensations" (Darwin, 1872, 33). Following Herbert Spencer (1863, 138), he uses the words emotion and sensation to refer to classes of feelings (Darwin, 1872, 33fn). For Darwin, who we can take as a major flag carrier for the common sense view James identifies, emotional feelings maintain an important causal role in emotional episodes. Conversely, for James, the emotional experience is much more obviously an effect than a cause.

1.2.2 “Perception of the exciting fact”: sensation, perception, and cognition

The central component of Ellsworth’s appraisalist interpretation of James is that term *cognition*. Ellsworth does not say what constitutes cognition; she merely deems Jamesian perception a cognitive act. In fact, she makes no mention of whether the appraisal process she holds James to imply is a special form of perception, or whether all perception (for James) involves appraisal. However, one can assume that James was aware of how he was using his own terms, so one can then assume that his definition of *perception* does not vary throughout the *Principles*. If Ellsworth is right that, for James, emotional reaction-triggering perceptions are cognitive appraisals, then all instances of perception are instances of cognitive appraisal.

To assess this conclusion we might answer this question: is Jamesian perception a cognitive act? Superficially, at least, the answer is yes: James himself describes perception as cognitive, as well as sensation (James, 1890, II: 1). We need not look past his writings on the emotions to see this statement: “Without the bodily states following on the perception [of the stimulus], the latter would be purely cognitive in form, pale, colorless, destitute of emotional warmth” (James, 1884, 190; 1890, II: 450; 1992, 352). So, it looks like Ellsworth has some solid textual evidence supporting her appraisalist interpretation.

But James’s use of the term *cognition* is very different than the variations employed in contemporary philosophy of mind and psychology. Indeed, appealing to James’s employment of *cognition* as a means of rendering his theory of emotions cognitivist is to be misled by the scent of a red herring. James builds his theory of perception in a way that undercuts the demands of appraisalism.

In his 1885 paper “On the Function of Cognition”, James argues that any feeling (or any thought) can be cognitive provided it resembles some external “reality”—a thing outside the mind—and helps the organism in acting towards that thing in some fashion. (In modern lingo, this makes James a representationalist of sorts.) He elaborates on this in the *Princi-*

ples, particularly in chapters 8 and 9 when discussing the distinctive functions of two forms of cognition: *knowledge of acquaintance* and *knowledge about*. After having described the distinction in terms of the degree to which a feeling suggests relations between itself and other feelings, he explains that some feeling is rendered *knowledge about* once the mind has “operated” on it, subjected it to some sort of analytic treatment (James, 1890, I: 221-222).

This spectrum between *knowledge of acquaintance* and *knowledge about* overlays James’s spectrum between sensation and perception. An examination of the latter spectrum elucidates the former. Sensation differs from perception “only in the extreme simplicity of its object or content”; it functions to provide *knowledge of acquaintance* while perception provides *knowledge about* (James, 1890, II: 1-2). But this distinction is of degree rather than kind:

The nearer the object cognized comes to being a simple quality like ‘hot,’ ‘cold,’ ‘red,’ ‘noise,’ ‘pain,’ apprehended irrelatively to other things, the more the state of mind approaches pure sensation. The fuller of relations the object is, on the contrary; the more it is something classed, located, measured, compared, assigned to a function, etc., etc.; the more unreservedly do we call the state of mind a perception, and the relatively smaller is the part in it which sensation plays. (James, 1890, II: 1)

Early in life we encounter many kinds of objects with many kinds of properties, and our experiences of them are not sophisticated. Thus, most of our earlier experiences will fall close to the sensation pole. But few will sit right at the edge of the spectrum; very rarely, even in early life, will we experience a “pure sensation”. Instead, as our brains form associations between different sensations and the contexts within which they are encountered, recurring sensations become increasingly rich in information: *knowledge about* emerges, and we cease to be epistemically blind as to the intentionality of our experiences. In adult life, we never experience pure sensations that convey merely *knowledge by acquaintance*. Instead, *every*

conscious experience is intentional and bears some information.

So, Jamesian perceptual states are indeed “cognitive” mental phenomena—as are those states we more readily call sensations. But this conclusion is superficial when viewed under the spotlight of appraisal theory. What is of relevance to an appraisalist interpretation is not whether James calls the components involved in triggering emotions “cognitive” (according to his definition), but whether he proposes that the bodily changes preceding emotions are triggered directly by internal psychological means (appraisals) rather than by external stimuli. What is relevant, that is, is whether there is any intermediate processing of meaning, on the part of the organism, between a stimulus and the bodily changes that follow.

It does not appear, for James, that perception fits the bill of intermediate mental processing. While perception provides knowledge about its objects rather than mere *knowledge of acquaintance*, it does so without any episodic inferential machinery.¹² Instead, *knowledge about* arises from associative mechanisms that build complexity into mental representations over time, through repeated experiences. The kinds of relations that an object of the mind comes to possess are those we have stumbled across in previous encounters with objects in the world. Since we perceive objects embedded in situations, the more kinds of situations in which we encounter a given object kind, the more opportunities we have to acknowledge relations between that object kind and other aspects of reality. Thus, when James says that *knowledge about* results from feelings being “operated” upon by the mind, he merely means that those feelings are situated within a network of related feelings by way of the brain’s associative mechanisms.

A perceptual representation can be informationally rich, providing plenty of *knowledge about*. For such a representation to become informationally rich, it must be trained through repetition over time. We come across objects within environmental contexts and those contexts link objects to other objects in particular ways. We experience those relationships and

¹²See James’s discussion, and rejection, of Helmholtz’s argument that perception involves unconscious inference (James, 1890, II: 111-113).

they are combined (by the laws of habit and association) with one another in the brain: familiarity with an object enriches its neural representation. Again, though, this occurs by way of association, not inference. The associative mechanisms underlying perception, in fact, do not get much of a psychological description. James instead describes association in terms of brain physiology (see, e.g., James, 1890, II: 111-113). When the sense organs are stimulated, there is a direct route of activity towards a region of the brain (presumably defined functionally rather than anatomically) that serves as a representation of what stimulated the sense organs. There is no inference made over the sensory inputs *in the moment* in order to trigger the representation; it is triggered directly as a result of sensory input. So, a perceptual representation results not from episodic mental work, but from gradual, environmentally informed, psychophysiological training. Therefore, Jamesian perception is epistemically direct even if it is representational at base.

How does this conclusion impact the appraisalist interpretation of James's theory of emotions? Recall that the crux of the matter is whether emotional reactions (i.e., bodily changes responsible for producing emotional experiences) are triggered by cognitive processing of sensory information. If the kind of perception that triggers emotional reactions is the kind of perception described above, then regardless of how near to either end of the sensation–perception spectrum it sits, that state need not be influenced by immediately prior mental processing in order to do its work. In James's words, "particular perceptions certainly do produce wide-spread bodily effects by a sort of immediate physical influence, antecedent to the arousal of an emotion or emotional idea" (James, 1992, 352-353, emphasis removed).

This is true regardless of how informationally rich and "full of relations" it is: for James, any stimulus-triggered representation counts as an intentional sensation or perception, provided its contents include some sort of interjection ("*lo! there! ecco! voilà!*"), article, or demonstrative pronoun (James, 1890, I: 222)—anything designating that the perceived object has a spatiotemporal index relative to the perceiver's body. Thus, when James talks about

the perceptions involved in triggering emotional reactions, these need not be informationally rich; the sensation or perception merely needs to be attuned to the present situation in a rudimentary fashion. As Gary Hatfield points out, a fear reaction can result from registration of a simple bear-shape or bear-motion rather than from perception of a detailed, multifaceted object (Hatfield, 2007, 420).

The appraisalist could push back here, objecting that even if sensations and perceptions immediately trigger bodily changes, James owes us an account of the mechanism by which this occurs. To this end James appeals to Darwinian adaptation: emotional reactions, *qua* instincts, result from innate associations between certain sense organ activities and sets of bodily changes (see James, 1890, II: 678-686; summarized in Hatfield, 2007, 417ff; see also James, 1890, I: ch. 2). This move has remained popular throughout the literature on emotions and instincts since James and it finds support from empirical investigations. For example, Luc Arnal and colleagues (Arnal et al., 2015) have shown that acoustic roughness (heard in natural alarm signals such as screams, and in artificial alarms such as smoke detectors) selectively activates the amygdala, which is heavily involved with producing fear responses.¹³ Relatedly, James points out that neonates are alarmed by any sudden sound, and only gradually learn to distinguish innocuous from suspicious sounds (James, 1890, II: 78fn).

Moreover, he explains that instincts are trainable by way of association. Much less ecologically obvious stimuli (such as awkward social interactions or wistful imaginings) can become associated with emotional reactions even if their downstream instinctual behaviors would provide little ecological benefit for the individual (James, 1890, II: ch. 24). The processes by which such associations are formed is complicated, and I can only briefly describe them here. They will typically involve deliberately attending to ideas and cognizing, say, their *social* relations, but the result will be the same kind of association as those formed due to our bodily ecological concerns (such as bear-fear associations): “A nervous tendency to

¹³Thanks to Michael Kaplan for pointing me towards this source.

discharge being once there, all sorts of unforeseen things may pull the trigger and let loose the effects. That among these things should be conventionalities of man's contriving is a matter of no psychological consequence whatever. The most important part of my environment is my fellow-man" (James, 1884, 195). Human minds, he continues, are extraordinarily sensitive to others' attitudes of them; we can perceive the intentions in one's gestures, looks, and actions (James, 1884, 195-196). Though we may require complex cognitive processing to associate socially determined stimuli with appropriate emotional reactions, those associations then can be triggered directly by perceiving such actions repeated thereafter. The episodic processes that cause a fear response to a bear or an anger response to an insult are causally analogous.

This point appears to have been missed by Laird and Bresler (1990, 645), who take themselves to "part ways" with James: while they provide empirical—though "indirect and inferential"—evidence that socially contingent situations can trigger emotions, the Jamesian picture (as they see it) relies on innate mechanisms only. In fact the evidence they cite is quite friendly towards James's actual claims. Damasio (1994, 129) is similarly mistaken in attributing to James an "inflexible" emotion-triggering mechanism. Clearly, certain innate stimulus–response pairings are more plastic than others.

Unfortunately, this response will not convince those who demand that emotional experiences themselves, and not merely their perceptual prerequisites, include an intentional aspect. Ellsworth's appraisalist revision of James initially appeals because its implied cognitive machinery makes it easier for emotions to be about their stimuli: we are scared *of* the spider, happy *that* our team won, angry *at* the person who insults us. But it too fails to imbue emotions proper with intentionality. Furthermore, although a given perception may be informationally rich, we are yet to see from James an explanation of the mechanisms by way of which perceptions lead to bodily changes. Appeal to innate and trainable associations between stimuli and responses can only provide us with reliable predictions about which

perceptions cause which changes—a powerful resource in itself—but such associations do not much inform a thorough psychological explanation of the intentionality of emotional experiences and the mind–world relationship.¹⁴ Nor do they help us to see satisfied the three intuitive desiderata of emotions mentioned on page 35. We are yet to defeat Deigh’s charge of epiphenomenalism.

1.2.3 Meeting the three intuitive desiderata

In a paper titled “William James on Emotion and Intentionality”, Matthew Ratcliffe argues that James rejects the distinction between affect (i.e., emotional experience) and cognition that is typically taken for granted (Ratcliffe, 2005, 179). This distinction is assumed by appraisal theorists, as shown in the quote from Lazarus on page 37 in which he distinguishes the broader category of cognition from specifically emotional cognitive appraisals. It also grounds criticisms of James. According to Ratcliffe, however, James claims that not only are cognitive states affective states; affect is also “inextricable from intentionality” (2005, 180). What he ends up with is a reading of James’s theory according to which one’s intentional perceptual content is dependent upon one’s emotional state: James’s view “incorporates a phenomenological conception of [intentionality], according to which the *experiential presentation* of things is essentially bodily and affective” (Ratcliffe, 2005, 184, original emphasis).

Ratcliffe’s claim relies on textual evidence from James’s later writings (from the early 1900s), after he has articulated his metaphysical theory of radical empiricism. Not every reader will be convinced that the earlier and later programs James builds are compatible in the way Ratcliffe’s interpretation presupposes. However, Ratcliffe does point out some important implications of James’s theory regarding emotion and perception. These features can also be found in the *Principles*, thereby sidestepping questions about weaving the later

¹⁴In fact, any theorist who proposes that perceptual or other cognitive states trigger emotional reactions owes such an explanation, unless we are happy to concede to the behaviorists’ pragmatic limitations, namely, that psychology is in the game of prediction and control of behavior (see Watson, 1930, 11). The Jamesian is not alone in her need to solve this puzzle.

James into his earlier works.

Ratcliffe's reading can be summarized as follows (Ratcliffe, 2005, 184-186). Human nervous systems, like those of other organisms, are evolutionarily pre-tuned to ecological niches: our perceptual systems are sensitive to certain features inherent in the kinds of objects that can impact our well-being. Thus, we respond selectively to environmental features that are relevant to our potential behaviors. This tuning is seen at the physiological level rather than at the level of perceptual content (i.e., what we might today call the "cognitive level"). In other words, it is not the features of the *object as perceived* that elicits a reaction, but stimulation of the physiological mechanisms (e.g., retinoneural pathways) that constitute perceptual systems. As patterns of excitation in sensory systems occur, they summon "good" or "bad" feelings. These feelings help to structure and motivate behavior by prompting us towards objects that feel "good" and away from objects that feel "bad". And since the "good" and "bad" feelings are triggered by objects occupying our ecological niches, the behaviors they structure systematically conform to our needs.

We should pause at this point; an objection looms. James's original claim is that affect is due to bodily changes, but Ratcliffe's reconstruction apparently has feelings leading to bodily changes, namely approach or avoidance behavior. And that looks like a position James would wish to avoid, given its similarity with the common sense theory of emotion. In fact this objection can be met quite easily. Recall James's rule that any change in the body is accompanied by a change in feeling. This rule must apply not only to emotional feelings, but to all feelings in general. Emotional reactions are largely due to efferent nerve signals, such as those that trigger muscular contractions across the body. Yet excitation of sensory systems by an external stimulus is just as legitimate a bodily change as is, say, an emotional reaction such as crying or laughing. More formally: bodily changes occur due to both efferent *and* afferent nervous signals. As James writes: "Any object, if *immediately* exciting, causes a reflex accommodation of the sense-organ, and this has two results—first, the object's increase in

clearness; and second, the feeling of activity in question. Both are sensations of an ‘afferent sort’ (James, 1890, I: 435, original emphasis).

With these bodily feelings causally accounted for by way of physiological stimulation rather than perceptual content, we avoid converting James’s theory into a version of the common sense view he explicitly sought to overturn. Nor must we concede to an appraisalist reading, since the means by which our bodily changes relate to stimuli is reflexive rather than cognitive.

The picture we have so far is, however, rather unsatisfying. We are yet to see how the microscopic feelings produced during afferent sensory stimulation contribute to the experiential presentation of perceptual objects. Ratcliffe’s answer is that the bodily changes felt during an act of perception contribute to the phenomenal structure of the percept by shaping the way things appear to us. The quotation from James at the end of the previous paragraph helps us make sense of this. The physiological act of perceiving—sequenced nervous activity between the sensory organs and the brain—triggers microscopic, reflex-like physiological responses. Because of this, Ratcliffe argues, any perceptual representation will have woven through it an affective reflection of our bodily concerns: “The experienced world is always a reflection of our capacities, needs and concerns. Emotions are indeed reports of bodily states, but, as the structure of experience is inextricably tied up with needs, concerns and bodily capabilities, such feelings are not wholly ‘internal’ to the subject but world-orienting and, indeed, ‘world-making’” (Ratcliffe, 2005, 190).

What does Ratcliffe mean by this term *world-making*? Once more, we come back to James’s notion of the self at play in his general psychological theory. He takes it as a “fundamental psychological fact” that a mind takes interest in anything that can impact it (James, 1890, I: 289). And, as we saw throughout section 1.1.2, the feeling of self is a product of bodily activity. Therefore, the physiological activity underlying perception is influenced not just by stimulus objects, but simultaneously by the current states of the body and the brain within

it. It is only logical that the corresponding mental representation of the perceived object should be *interesting* to the individual, given its content involves some quality or report of the “capacities, needs and concerns” of the embodied self. As James explains:

... every creature has a certain selective interest in certain portions of the world, and... this interest is as often connate as acquired. Our *interest in things* means the attention and emotion which the thought of them will excite, and the actions which their presence will evoke. Thus every species is particularly interested in its own prey or food, its own enemies, its own sexual mates, and its own young. These things fascinate by their intrinsic power to do so; they are cared for for their own sakes. (James, 1890, I: 320, original emphasis)

The intentionality of a perceptual state thus depends upon the intentionality of an emotion that is inextricably bound up with that episode of perceptual processing. It is not that an object is perceived and *then* assessed for its impact upon the individual; rather, the object is perceived *as related to the individual*, since the concerns of the self guide the perceptual capacities of the organism to begin with. Contrary to the demands of the appraisalists, a Jamesian emotion need not rely upon prior transformation of sensory information into meaningful, evaluative judgments. Instead, Jamesian emotional experiences, as intentional states, help to determine the intentional content of perceptual states. When an organism's sensory system produces a pattern of activity correlating with ecologically interesting things, this activity triggers microscopic emotional reactions that present the object registered through the senses as something that can impact the organism. In short, the emotions themselves do most of the work that appraisalists claim to be prerequisite for them to occur. James's theory thus satisfies the three intuitive desiderata while maintaining its noncognitivist status. Emotions neither reduce to nor causally depend upon judgments or appraisals.

1.2.4 Micro-emotions and macro-emotions

The previous subsection gives us the following account of an episode of Jamesian emotion:

1. First, a stimulus object stimulates the organism's sensory organs (say, a bear-shadow projected on to the retinae).
2. Activity of the organs, and of the nerves leading from them to the brain, produces obscurely felt sensations that reflect their sources' physiological status, thereby reflecting the organism's natural bodily concerns and contributing to the content of the perceptual representation of the stimulus.
3. If the sensory stimulation is part of an associative pathway that correlates with the organism's ecological interests, that pathway will lead to further and more severe activity across the body, constituting the emotional reaction.
4. The emotional reaction is then felt acutely as an emotional experience.

Note that step 2 displays the same relation between physiological activity and phenomenal experience as does the now-familiar relation between emotional reactions and emotional experiences. In other words, activation of sensory systems produces experiential states that can legitimately be categorized as emotional experiences; these occur simultaneously with, and partly constitute, perceptual representations. Therefore, on James's theory properly reconstructed, emotional states are omnipresent in consciousness. This claim is consistent with his conception of the self, as well as with his claims about the "intellectual" emotions—those that are experienced without an explicit bodily expression—in the second half of his 1884 article and his chapter on emotions in the *Principles*.

However, there is yet room to raise an issue with this reconstruction. Even if feelings of microscopic bodily changes during afferent sensory stimulation (mentioned in step 2 above) are in themselves emotional, properly speaking, they are certainly not experienced *as* emo-

tions (unlike those mentioned in step 4). Indeed, we seem perfectly able to experience perceptual states that do not carry with them any obvious emotional quality. And the idea that perceptual states are disinterested (thus unemotional) by default is what motivates appraisal theorists such as Arnold and Lazarus to develop their positions. Generally, emotional experiences are of a much more vividly felt quality than the microscopic feelings at work in perceptual processes.

We can characterize the distinction of importance here as that between *micro-emotions* and *macro-emotions*. While micro-emotions are not felt *as* emotions—they are felt as elements of perceptual processes, and typically felt obscurely rather than acutely—macro-emotions are those states we distinctively feel as emotional experiences. Such macro-emotions include the “standard” or “coarser” emotions described in section 1.1.1. How do macro-emotions come to be felt as emotional experiences, if micro-emotions do not? Call this the *problem of phenomenological scale*.

An answer lies in the kinds and degrees of bodily changes occurring in either class. The bodily changes underlying micro-emotions are those involved in accommodation of the sensory systems. We saw this in the quotation from James on page 49: objects that excite the sensory organs cause those organs to accommodate so to gather more sensory information from those objects. A clear example is ocular accommodation: if you voluntarily defocus your eyes while gazing at an object, and then ‘let go,’ so to speak, of your ocular muscles (specifically the ciliary muscles and the pupils), you will notice your eyes ‘snap’ into focus on the object, increasing its clarity in your visual field. This is an automatic, reflexive activity on the part of the body that relies on stimulation of the retinae. (If you struggle to defocus your eyes, you can see the same result by playing with the autofocus function on your camera.)

The bodily changes involved in macro-emotions can be thought of as accommodation on a much larger scale, distributed and coordinated across the body. Such changes will be triggered by the patterns of changes determined during the physiological act of perceiving

the stimulus (such as retinal stimulation and ocular accommodation). These might include changes in orientation of the body (or some of its sensory organs), say, towards the stimulus so to gather more sensory information about it—a widening of the eyes, a turning of the head, and so on. Or the body may be turned away or shielded so to minimize harmful impact, such as we see when a finger is quickly thrust towards the eyes. And the effects on the internal parts of the body may include a quickening of the pulse, release of adrenaline into the blood, and a general tensing of the skeletal muscles, thereby constituting the characteristic emotional reaction of fear. As we saw above in section 1.1.2.2, these bodily changes prepare the organism to carry out some instinctual behavior in response to the ecological significance of the stimulus. It is such downstream bodily changes as these that involve activity of, and coordination between, more of the body's muscles and organs (and larger ones at that). Wholesale changes across the body bring wholesale, immediate differences in the body's "sounding-board", as James puts it (1884, 202), such that larger-scale changes underlying macro-emotions are felt *as* emotions themselves.

This solution to the problem of phenomenological scale raises two secondary problems. The first is the *problem of evaluative scale*: even if we concede that micro-emotions convey 'good' or 'bad' feelings in response to stimuli, this does not imply that macro-emotions too possess any sort of evaluative quality. Julien Deonna and Fabrice Teroni address this problem in their recent defense of Jamesian emotion theory, arguing that emotional experiences are states of felt action readiness (Deonna & Teroni, 2017, 59). To summarize their argument (2017, 61-62): When we are experiencing an emotion, our attention is directed outward, at the environment. When this is the case, the body feels differently than when we are introspecting. Thus, states of action readiness "appear in consciousness as ways of being engaged with the environment" (2017, 61). So, states of action readiness are outward-looking; emotions are experienced as one's potential reaction to the ecologically impactful features of one's environment. In this sense, Deonna and Teroni's solution to the problem of evaluative scale

(or what they call the “Objection From Evaluation”) helps to emphasize the personalist current running through James’s general psychological theory. Emotional experiences involve a felt self as one participant in a two-sided relationship.

The other issue is the *problem of restricted stimuli*: how can we reconcile the idea that emotions originate in sensory accommodation with the plainly obvious fact that emotions can be triggered by objects of thought? It’s very possible to feel anxious at the idea of losing one’s job. But there is arguably no *perceivable* stimulus available to trigger such an emotional experience. James’s view as presented appears to ignore that emotional stimuli can be triggered by a wide range of phenomena. Although we saw section 1.2.2 that the range of stimuli to which an organism can become emotionally sensitive will widen through associative learning, couching micro-emotions in sensory organ accommodations restricts the range of bodily causes of macro-emotions to those involved in sensory perception. Thus, emotions cannot be triggered directly by thoughts. And this looks contrary to James who allows for emotions to be triggered by “the passage of the exciting train of ideas” (James, 1884, 189).

James himself can provide the solution to this problem. In the opening chapter of the *Principles* he gives us a rule regarding the mind-body relationship. This is best presented verbatim:

Mental phenomena are not only conditioned *a parte ante* by bodily processes; but they lead to them *a parte post*. That they lead to *acts* is of course the most familiar of truths, but I do not merely mean acts in the sense of voluntary and deliberate muscular performances. Mental states occasion also changes in the calibre of blood-vessels, or alteration in the heart-beats, or processes more subtle still, in glands and viscera. If these are taken into account, as well as acts which follow at some *remote period* because the mental state was once there, it will be safe to lay down the general law that *no mental modification ever oc-*

curs which is not accompanied or followed by a bodily change. (James, 1890, I: 5, original emphasis)

And later in the same volume he writes more specifically of thoughts about objects:

...there are good grounds for believing that even intellectual attention, attention to the *idea* of a sensible object, is also accompanied with some degree of excitement of the sense-organs to which the object appeals. (James, 1890, I: 434, original emphasis)

Thus, provided that ideas trigger stimulation—even if very faint and microscopic—in some parts of the sensory systems, micro-emotions will occur, which may eventually lead to macro-emotions in response to (and about) nonpresent stimuli.

Throughout this section I have offered a reading of James that renders Jamesian emotions as intentional, information-bearing, and motivational states. Not only is Ellsworth's appraisalist interpretation of James incorrect; it is also an unnecessary augmentation of the theory that, of its own resources, can satisfy the three intuitive desiderata. Cognitivist revisions of James's theory of emotions are unnecessary in general, because James is innocent of the charge brought against him: that he takes emotions to be epiphenomenal, unintelligent feelings. A major outcome of the reconstruction offered in this section is that we need not accept the conclusion ((C2) in the introduction) that we must tease out of James's theory implicit cognitive machinery in order to render it compatible with today's cognitivist theories of emotion. Instead, James's theory of emotion is informative in large part because it can stand independently of, even in contention with, cognitivism.

Conclusion

The aim of this chapter has been to demonstrate that James's theory of emotions is commonly misunderstood by contemporary commentators. These misconceptions invite pre-

mature conclusions regarding the theory's relevance today. They should be eliminated in favor of a faithful portrayal. Understanding the theory requires contextualizing it: namely, recognizing that James employed an unorthodox terminology throughout his psychological framework. Once one properly appreciates that, for James, emotions are nothing more than sensations (now properly cashed out in Jamesian terms) caused by bodily changes (section 1.1), and that those bodily changes are caused not by internal processing of stimuli but by direct, reflexive sensory-neural systems (section 1.2), one can grasp the most important contribution of his theory. That is the idea that we need not appeal to a complicated cognitive architecture in order to account for an organism's sophisticated emotional reactions to the world and its corresponding experiences.

Of course, our emotional tendencies mature as we individuals develop, and cognition will undoubtedly play a formative role in shaping our emotional lives. But that is not to say that we need cognitive processes that operate over sensory information in order to react emotionally to a situation. An emotion is a bodily sensation, and its triggering conditions are not different in kind or psychologically unique compared to other means of responding to the world. Since emotions' triggers are reflexive and automatic, James's theory can extend out of human psychology and across a broader range of organisms regardless of their cognitive limitations. That is, it helps to explain how cognitively simple creatures such as human infants and nonlinguistic animals—for whom the world is experienced primarily as “one great blooming, buzzing confusion” (James, 1890, I: 488)—can experience emotions. On the other hand, as those triggering conditions (in humans at least) are trainable and developmentally plastic, we can describe why a person can feel similarly emotional during drastically different situations, and why two individuals can feel differently under similar circumstances. James's theory is thus doubly fecund: it spreads across species, and it scales across situations of greater or lesser cognitive complexity. That promise makes his theory well worth comprehending.

A major way in which a serious treatment of James's theory of emotion can benefit us is in investigating the interactions between emotion and perception. The general idea that emotional experiences, or affects, can influence perceptual states is not exclusive to Jamesian theories. In an article in *Nature Neuroscience*, Ralph Adolphs describes a neuropsychological study that provides evidence in favor of the hypothesis that "emotion can directly influence sensory processing, and can do so at surprisingly early stages" (Adolphs, 2004, 1167). Lisa Feldman Barrett and Moshe Bar argue for the same hypothesis from distinct empirical evidence, claiming that "the brain's ability to see in the present incorporates a representation of the affective impact of those visual sensations in the past" (Barrett & Bar, 2009, 1325). Relatedly, Simona Ginsburg and Eva Jablonka appeal explicitly to Jamesian emotion theory in an article on the evolutionary nature of experience (Ginsburg & Jablonka, 2010a),¹⁵ a topic that will be explored in Chapter 3.

Properly understood, James's theory of emotion has the potential to influence contemporary theorists looking to develop fruitful approaches to studying the human mind and the world to which it reacts. It offers more than mere reflexivity, while relying on less sophisticated psychological machinery than most researchers think is necessary. Based on the above, there is simply no need to replace a mentally simplistic Scarecrow with a hyperintellectual Tin Man, for James is neither.

¹⁵Thanks to Peter Godfrey-Smith for informing me of this source.

Chapter 2

James's legacy: alternatives and alterations

Introduction

The previous chapter introduced and analyzed William James's theory of emotion. It also examined how the theory has been received and interpreted—and misinterpreted—since its initial publication in 1884. While the intent was to defend James's theory against naïve dismissals and inaccurate presentist reconstructions, a third trend went unexamined: the idea that the theory can be made compatible if we tweak it a bit and supplement it with psychological mechanisms identified in our current scientific theories.

This chapter picks up this thread and examines contemporary theories of emotion that sprouted from Jamesian seeds. Mainstream psychology in the US has seen plenty of growth and conceptual change over the past century: the personalist, introspectionist strand of James's and others' psychological approaches gave way to various species of behaviorism in the 1910s and '20s; a pattern of paradigm shifts brought mental representations, “internal” cognitive mechanisms, and the influence of social factors on thought and behavior into

focus around the 1960s and '70s; and the development of brain imaging technologies facilitated the rise of social, affective, and (most prominently) cognitive neuroscience from around the mid-1980s. The massive conceptual and methodological shifts in the general domain of psychology mean James's theory in its original form does not neatly fit into an extant ontology of the mind. Hence, those researchers who take James's theory seriously today endeavor to revisit and revise its elements using the tools of social and cognitive psychology, neuroscience, and updated philosophies of mind.

Three theories stand out as developments of a broadly Jamesian approach to explaining emotion. Each is examined below. Before they are discussed, however, it is important to see what the Jamesian program contends with. Therefore, the chapter opens with a discussion of cognitivist theories of emotion and the problems they face (section 2.1). Chief issues for cognitivist theories are their hyperintellectual demands: if emotions depend, as cognitivist theories claim, on complex, self-referencing or conceptual cognitive machinery to make them manifest, then these phenomena are inaccessible to creatures other than humans. Although each, or at least some, of these problems may turn out to be surmountable, the aim here is to show that cognitivism is not straightforwardly a preferable alternative to Jamesian theories.

The two sections following that introduce and analyze the three aforementioned theories inspired by James. Section 2.2 reviews an influential attempt to update James's theory in light of the rise of social psychology during the years either side of World War II: a kind of marriage of cognitivist and Jamesian principles developed to account for emotions as intrinsically interpersonal phenomena. Unfortunately, this cognitivist alteration of the Jamesian theory is insufficient in scope to explain the ranges of emotion expressed and experienced by most creatures—that is, the cognitivist alteration brings with it the cognitivists' headaches. Section 2.3 then outlines and scrutinizes two so-called neo-Jamesian theories of emotion: theories which purport to be faithful both to the spirit of James and to the progresses of

recent psychology, neuroscience, and philosophy.

Attempts to progress the Jamesian program should be celebrated. But these three theories all share an assumption that the previous chapter worked to dispel: namely, the assumption that emotions depend causally on other psychological processes such as cognition and perception. Section 2.4 questions this assumption: Do emotions require mental causes? This final section offers a brief overview of the argument that will form the basis of Chapter 3. According to this argument, emotions are logically prior to other kinds of mental states. Therefore, emotions do not require mental causes.

2.1 Problems with cognitivism

One might wonder what value is to be gained in pursuing a noncognitivist theory of emotion such as James's. The dominant position in contemporary emotion research, cognitivism, asserts that emotions are, or are the results of, certain operations of an organism's cognitive machinery (judgments or appraisals). While few authors offer solid definitions of cognition, a general theme in the camp sees the term pick out any processes whose outputs are representational and meaningful in nature and not merely reflexive. Emotions then are the products of cognitive processing of some kind of information, whether perceived, conceived, or imagined.

In advocating the idea that an organism's emotions are necessarily dependent upon its cognitive capacities, the cognitivist faces three major challenges. Firstly, by taking humans' experiences as prototypical of emotions, cognitivist theories are biologically narrow in scope (2.1.1). Secondly, by reducing emotion to evaluative judgments, cognitivists struggle to accommodate feeling within their explananda, despite the demands of common intuition (2.1.2). Finally, the cognitivist's preferred solution to this second problem commits her to a messy conception of *an emotion* as a cluster of psychophysical events rather than a single

psychological phenomenon (2.1.3).

2.1.1 Cognitivism is biologically narrow

Cognitivism is the dominant approach throughout the cognitive sciences, and it generates many fruitful lines of investigation into human emotionality. But cognitivism is restrictively anthropocentric: it takes as its paradigm cases the various emotions described by humans, often at the expense of other, psychologically simpler creatures. The cognitivist takes for granted that creatures whose neural and cognitive architectures do not support cognitive evaluation of exteroceptive stimuli, or of mental representations, cannot experience emotions (e.g., Lazarus, 1984). And those animals that can experience them will manifest them through far less complex and symbolic cognitions than will humans (Lazarus, 1982). Richard Lazarus is also skeptical about the potential of investigating animal cognition at all:

... the methodological difficulties of evaluating cognitive activity in infrahuman animals should make us wary of accepting statements about what animals can or cannot accomplish cognitively. I am doubtful that any statement about the absence of cognitive appraisal processes in an animal, even a simple evaluative perception studied experimentally or in the field, could be made without substantial doubt. (Lazarus, 1984, 127)

By extension of cognitivist logic, then, we should assume that (some) non-human animals possess cognitive appraisal mechanisms and therefore can experience emotions. Perhaps a Fodorian language of thought is available to emotional animals.¹⁶

The cognitivist's other option is to propose a conceptual split between two senses of emotion: those experiences had by concept-lacking organisms that we call "emotions" are at best some sort of rudimentary affective state; whereas only more sophisticated minds, such as

¹⁶Thanks to Gary Hatfield for this point.

those of humans, can experience emotions in an intellectually tractable or “philosophically rich” sense. Lisa Feldman Barrett—not quite a cognitivist, but whose constructivist theory of emotion demands the operation of emotion concepts—argues that discussions of animal emotionality cannot resemble discussions of human emotionality due to humankind’s exclusive access to emotion concepts (Barrett, 2017). Humans have human emotions, animals have animal emotions, and never the twain shall meet. (Except maybe in some human-raised great apes who learn human languages. There the twain may meet.)

Consequently, researchers across philosophy and the cognitive sciences favor the latter conception: theories of emotion deal with human phenomena exclusively. Since cognitivism defines emotions in terms of evaluative judgments, emotions are theoretically reduced to (that is, redescribed as) variations or subspecies of the phenomena that the information-processing models of the cognitive sciences are built to examine and explain. Too bad for any creature who reacts to the world through exclusively reflexive or nonconceptual means: if the cognitive sciences cannot describe its experiences, it is expelled from the scientific realm of emotionality.

Even within the human domain, this commitment to a narrow conception raises problems. Certain human emotions, such as shame and guilt, intimately connect to normative judgments and therefore appear well placed for explanations using cognitivist resources. However, some of our more urgent and spontaneous states—fear and startle responses, for instance—evade such approaches. Robert Solomon, for one, is simply not interested in accounting for the reflex-like responses that characterize such “emotions”. He writes:

... I am interested, to put it polemically, in processes that last more than five minutes and have the potential to last five hours, five days, or five weeks, months, or even years. I am interested, in other words, not in those brief ‘irruptive’ disturbances but in the long-term narratives of Othello, Iago, Lily Bart and those of my less drama-ridden but nevertheless very emotional friends. I am interested

in the meanings of life, not short-term neurological arousal. (Solomon, 2003, 2)¹⁷

Peter Goldie (2000) similarly defines emotions as complex, long-lasting states with narrative structures, though he is also interested in accounting for the “emotional episodes” that contribute feelings to such narratives.

That said, the idea that emotions are temporally extended like this is not a central tenet of cognitivism. The more pressing issue regards the ways in which emotions are triggered. This often leads theorists to another split between notions of emotion, with varying exclusionary attitudes towards the less sophisticated phenomena. Solomon’s most important claim is that “short-term neurological arousal” does not satisfy the triggering conditions for an emotion (*qua* phenomenon of theoretical interest). Yet there is little consensus on what might be the minimal triggering conditions for emotions. For instance, there has been much ink spilled in the debate over whether startle qualifies as an emotion due to its seemingly direct, reflex-like manifestations (see, e.g., Landis & Hunt, 1939; Ekman et al., 1985; Robinson, 1995). The underlying issue of whether emotion designates a unified class of phenomena can, in principle, be informed by the criteria determined by cognitivist theories. In practice, however, drawing such boundaries even within human psychology has proven a burdensome task (see, e.g., Cowie & Cornelius, 2003, and Chapter 4, section 4.3.1).

By contrast, noncognitivist theories can deal with the more urgent, unsophisticated phenomena (such as fear and startle) fairly well, allowing for them to qualify as legitimate emotions alongside guilt, shame, and the like. Where noncognitivist theories typically struggle, however, is in accounting for those socially influenced emotions (see also Damasio, 1994,

¹⁷Solomon finds himself in very good company among emotion theorists who turn to Shakespeare for inspiration. A quick search through my personal library of emotion literature finds Shakespeare’s works referred to in the following publications: Darwin (1872); Damasio (1994, 1999); de Sousa (2017); Deigh (2004); Ekman (2003); Ellsworth (2013); Evans (2002); Goldie (1999, 2000, 2002, 2004); Gottfried & Jow (2003); Griffiths (1997); Irons (1894); Johnson-Laird & Oatley (2008); Kieran (2010); Larsen et al. (2008); Lazarus (1991); LeDoux (2002b); Neu (2010); Nussbaum (2001); Ortony et al. (1988); Robinson (2005); Scarantino (2010); Solomon (1976). It’s unlikely this list is exhaustive.

130-131; Griffiths, 1997). Guilt and shame require access to social and moral norms: they are triggered by perceived deviations from contingent expectations. However, one should bear in mind that noncognitivism about emotion does not imply *anticognitivism* about emotion. An anticognitivist position denies that emotions ever are triggered through cognitive means (such as judgments of deviation from expectations). A noncognitivist position, to repeat, simply denies the *necessity* of the principle that all emotions are caused by cognitive processes. Thus, noncognitivist positions can potentially allow for a significantly wider set of triggering conditions for emotions than can either cognitivist or anticognitivist accounts. Psychological continuity therefore can exist across numerous species: my dog and I might be alike in our experiences of fear or elation. How far this continuity spreads can be determined only by looking at individual species. Whether a given species has emotional capacities is better determined through empirical investigation than through *a priori* reasoning.

2.1.2 Where do feelings fit in?

A quote from Martha Nussbaum exhibits the strongest demands of cognitivism: "... it makes sense to imagine that a thinking being, whether realized in matter or not, could care deeply about something in the world, and have the thoughts and intentions associated with such attachments. And that's all we really require for emotion" (Nussbaum, 2001, 60). Nussbaum's theory is a particularly strong version of cognitivism, but her fundamental claim echoes throughout the camp: emotions are or are caused by the organism's rich intentional mental states. Assuming emotions to be so reducible, researchers turn their attention to the puzzle of *differentiating* or *characterizing* emotion as somehow unique among cognitive phenomena. While emotion may reduce to cognition, it remains a discrete subspecies due to certain properties—properties whose identities are, to date, opaque to investigation and difficult to track through formal analyses.

If prompted, most people (including off-the-clock philosophers) will probably acknowl-

edge their present emotional states by attending to some *feeling*. Could this be an essential property that demarcates emotions from other cognitive states? Not necessarily. For Nussbaum, even “feelings” qualify as identity conditions for emotions only provided those states carry “rich intentional content – feelings of the emptiness of one’s life without a certain person, feelings of unhappy love for that person, and so forth” (Nussbaum, 2001, 60). Understanding “feelings” in this way, Nussbaum continues, leaves us without a way to differentiate them from the evaluative judgments themselves (Nussbaum, 2001, 60). Furthermore, other, nonintentional bodily feelings that frequently accompany emotions (such as feelings of muscle tensions and increased heart rates) are not conceptual components of emotions.¹⁸ As is hinted above, Nussbaum is happy with the idea that a bodiless, feelingless organism can nonetheless entertain legitimate emotions. If we follow Nussbaum’s construal, we can characterize a state as an emotion absent any phenomenality whatsoever; in fact, emotions can even occur unconsciously (Nussbaum, 2001, 60–61). I can be angry at someone without experiencing any psychic heat *and* without accessing through metacognition the judgment that constitutes my anger (that is, without *consciously* entertaining the propositional attitude *I am angry at that person*).

Since Nussbaum’s theory denies that feelings are necessary features of emotions, the cognitivist is faced with another conceptual commitment: either (i) emotions are not intrinsically related to feelings; or (ii) a token of the term *emotion* denotes not a single phenomenon, but a cluster of causally and temporally linked psychological and physiological occurrences.¹⁹ Nussbaum explicitly argues for the first option: *emotion* reduces to a form of value judgment. Let’s look now at the alternative for cognitivism.

¹⁸Goldie (2000; 2002; 2009) also theorizes that the kinds of feelings involved in emotions are intentional; he calls these *feelings towards*, distinguishing them from mere bodily feelings.

¹⁹In fact, these two options are not mutually exclusive. However, treating them as such allows the cognitivist to maintain the centrality of evaluative judgments in her conception of *emotion*.

2.1.3 Emotion as a cluster concept

Many, if not most, cognitivists today prefer the second option. While Solomon, for instance, proposes a view similar to Nussbaum's in his 1976 book *The Passions*, he has since become increasingly concerned with the role of the body and feelings in emotions (Solomon, 2003, 2004). Nonetheless, his updated theory remains cognitivist in nature despite his dislike of the label. While I take it as a given that any theory of emotion must account for the way in which we experience our episodes of fear, joy, guilt, and the like, I am remain unconvinced that cognitivist theories are well placed to satisfy that demand.

A cluster concept of emotion might deem a sequence of events an *emotional episode* (or just *an emotion*) if it involves certain necessary and sufficient constituents. Phoebe Ellsworth and Klaus Scherer describe a commonly proposed cluster concept: “[One] view, held by many appraisal theorists, is that appraisals are components of emotions—that the subjective experience of fear, for example, is the feeling of high attention, negative valence, high uncertainty about what is happening or one's ability to cope with it, and so on (in addition to the physiological and motor reactions elicited by these appraisals)” (Ellsworth & Scherer, 2003, 575). Moreover, an emotional episode like this may (in fact probably will) involve numerous reappraisals: an emotional episode can be indefinitely long and can undergo various metamorphoses. “Of course,” the authors continue, “when all the requisite appraisals occur, what the person feels is fear, not a collection of identifiable elements” (Ellsworth & Scherer, 2003, 575). On the cluster concept view, an emotion is a sum of its parts, but only some of those parts may be experienced.

Jesse Prinz asks the following question of cluster concept theories: “Typical emotion episodes...contain a number of components. There are thoughts, bodily changes, action tendencies, modulations of mental processes such as attention, and conscious feelings. But which of these things *is* the emotion? ... Can any given part be subtracted without losing the emotion, or are some parts essential” (Prinz, 2004, 3-4)? This is a dilemma for cluster

concept advocates. Prinz calls it the *Problem of Parts*.

If we cannot answer the Problem of Parts—if we cannot isolate that one essential component of our cluster concept that makes it an emotion (or makes the whole episode emotional)—then perhaps all of the components are essential. Prinz calls this counterpoint to the Problem of Parts, the *Problem of Plenty*:

The Problem of Parts asks: What components of an emotion episode are really essential to its being an instance of some particular emotion? The tempting answer is that all parts are essential. The Problem of Plenty then asks: If all parts are essential, how do they hang together together into a coherent whole? Put differently, the Problem of Parts asks for essential components, and the Problem of Plenty asks for an essential function of emotions in virtue of which they may have several essential components. (Prinz, 2004, 18).

Again we see a lack of consensus in the cognitivist literature. For every Martha Nussbaum who believes emotions need not include conscious experiences, there is a Magda Arnold who claims that emotions are felt action tendencies that must have evaluative appraisals as their direct causes (Arnold, 1960a). There is simply too much incongruity throughout the cognitivist camp to settle this dilemma by appeal to consensus. That is not at all to imply that the cognitivist cannot solve the Problem of Plenty, but it does present a conceptual roadblock to empirical investigation of emotion. Without settling on the set of phenomena to look for, we will continue to struggle to find anything concrete.

Conversely, Jamesian theories define an emotion as a simple, unitary mental phenomenon. An emotion is a feeling, with specific qualities, that functions in specific ways with respect to an organism's ecological concerns. It can have various causes and it can vary in its temporal extension. By identifying a single type of entity as emotion, the Jamesian approach avoids the analytic issues Prinz brings to attention. (Indeed, Prinz's own view, as we'll see, derives from broadly Jamesian inspirations.)

But James's program in its original form lacks precision and detail. The previous chapter presented James's theory of emotion as an appropriate guide to contemporary emotion research. The next two sections show some more recent efforts to progress the Jamesian program and build towards comprehensive theories of emotion.

2.2 An unsuccessful union: cognition-arousal theory

One of the more celebrated attempts to build upon James's insights is Stanley Schachter and Jerome Singer's (1962) study on cognitive and physiological cues. Schachter and Singer predicted that a state of physiological arousal, while necessary, is on its own insufficient to induce an emotion; a cognitive interpretation of one's physiological state, along with other factors such as one's circumstance and a dearth of alternative explanations, is also needed. They hypothesized that (i) if an individual is in a state of physiological arousal and he²⁰ has no immediate explanation for it, he will label and describe it in terms of cognitions available to him; (ii) if an alternative explanation for physiological arousal is available (e.g., injection of a drug), he will be unlikely to label it in emotional terms; and (iii) given the same circumstances, he will deem himself emotional only to the extent that he experiences physiological arousal (Schachter & Singer, 1962, 381-382).

To test their hypotheses, the authors designed a suite of scenarios to induce test subjects to experience either anger or euphoria. The experimenters recruited 184 male college students and told them they'd be injected with a vitamin that may affect their vision. In fact, the experimenters injected half of them with epinephrine (also known as adrenaline), the other half with a saline placebo. Epinephrine, of course, has an arousing effect across the body: "Shortly after injection systolic blood pressure increases markedly, heart rate increases somewhat, cutaneous blood flow decreases, while muscle and cerebral blood flow

²⁰While I typically designate "she" as the generic pronoun, the use of "he" is more appropriate here since the experimenters tested males exclusively.

increase, blood sugar and lactic acid concentration increase, and respiration rate increases slightly. As far as the subject is concerned the major subjective symptoms are palpitation, tremor, and sometimes a feeling of flushing and accelerated feeling” (Schachter & Singer, 1962, 382). Saline provides none of these effects.

These two groups were further divided into smaller groups and variously informed, misinformed, or left ignorant of the potential physiological effects of their injections. Each subject was then placed in a room for 20 minutes with a stooge who would behave in increasingly euphoric or angry ways. Subjects in the euphoric condition were engaged by the stooge to play some ad hoc games. (Say, scrunching bits of paper into makeshift basketballs to play hoops, or constructing paper airplanes.) Those paired with the anger-condition stooge were asked to complete an inappropriate questionnaire. (Questions ranged from slightly irritating—“What is your father’s average annual income?”—to highly unprofessional by 1960s standards—“With how many men (other than your father) has your mother had extramarital relationships?”) Following the sessions with their stooges, all subjects were asked to complete a questionnaire that would betray their self-reports of mood and physical condition.

The authors’ interest in this case was the subjects’ interpretations of their subjective physiological symptoms in light of their experiences with the stooges and scenarios. If indeed physiological arousal is insufficient to induce particular emotions, then epinephrine-injected subjects in both the anger and euphoria conditions should interpret the same set of drug-caused physiological arousal as symptoms of either euphoria or anger, depending on context. (That is, provided they didn’t have an alternative explanation for these symptoms, such as that they were typical side effects of the vitamin injection.)

Results of their study appeared to concord with their hypotheses. Placebo subjects reported themselves less euphoric or angry than did uninformed epinephrine-injected subjects (i.e., subjects who did not expect side effects). Subjects who had been told to expect

side effects (from the “vitamin”) reported themselves less emotional again. The authors consequently present these results as strong evidence in favor of the *cognition-arousal theory* of emotion: “emotional states may be considered a function of a state of physiological arousal and of a cognition appropriate to this state of arousal” (Schachter & Singer, 1962, 398).²¹

What do these findings imply for Jamesian theories of emotion?

2.2.1 Implications for the Jamesian program

Note that Schachter and Singer’s theory grates against two key tenets of the Jamesian program: the *specificity* and *constitution* of emotion.

James’s *specificity* claim holds that different types of emotional experiences “are associated with specific types of bodily feelings” (Deonna & Teroni, 2017, 55). That is, for a given emotion episode, the extents and locations of bodily change will exhaustively determine the phenomenal character of that episode. Conversely, Schachter and Singer’s experiment design reflects their assumption that differences in patterns of physiological arousal “are at best rather subtle and that the variety of emotion, mood, and feeling states are by no means matched by an equal variety of visceral patterns” (Schachter & Singer, 1962, 380).²² If indeed different instances of emotion are underspecified, or not at all specified, by different patterns of bodily arousal, then Jamesian theories appears unable to help us distinguish between two seemingly distinct emotional experiences. Other factors (most likely cognitive ones) then must be involved in the specification of emotion.

The *constitution* claim is that emotional experiences “are constituted by patterns of bodily feelings” (Deonna & Teroni, 2017, 55). This claim is logically distinct from the *specificity* claim: even if the latter is false, and there is no phenomenal distinction between two emotional experiences caused by two distinct patterns of physiological arousal (as Schachter and

²¹See Reisenzein (2017) for overviews of varieties of the cognition-arousal theory.

²²Let us ignore, for now, that Schachter and Singer erroneously constrain James’s physiological account to visceral activity. See Chapter 1 for more on this.

Singer allow), a theorist could hold that those two phenomenally identical states nonetheless constitute emotion in its entirety. That is, one could identify only a single phenomenal character for all (pre-reflective) instances of emotion, with any acknowledgement of variation accounted for by post-experiential self-reporting. But then the formal concept of *emotion* would lose much of its explanatory value: it would lack correspondence with historical and lay understandings of emotion that motivate the scientific project of explaining them; and it would be stripped of much of its normative force (such as variation of contributions to behavioral impulses). In short, denying the *specificity* claim leads to the redundancy of the *constitution* claim. Schachter and Singer's theory avoids so diminishing the concept of *emotion* by denying the *constitution* claim, instead positing a cluster concept: the term *emotion* refers to a function of both detailed cognitive and vague physiological cues. Robert Gordon points out that, although Schachter had, in earlier work, questioned whether emotions should or should not be equated with self-reports, "by the time of the Schachter–Singer study such doubts have been resolved, forgotten, or suppressed" (Gordon, 1978, 129). He characterizes their assumption: "If a subject reports that he feels very angry, then it is assumed that he is undergoing an emotional experience, and one that is distinct from that of fear or joy. And if by manipulating the subject's cognitions you can influence his verbal report of emotion, then it is assumed you are also influencing what he feels or experiences" (Gordon, 1978, 129).

Jamesian theories subscribe to both the *specificity* and *constitution* claims. Advocates must respond to the challenges raised by Schachter and Singer. The next two subsections defend the *specificity* and *constitution* claims against the Schachter-Singer experiment (2.2.2) and theory (2.2.3).

2.2.2 Issues with Schachter and Singer's experiment

Schachter and Singer's study has influenced a huge portion of the scholarship on emotion. It has also attracted criticisms. Some of them are directed at its treatments of physiological contributions to emotion self-reports. Rainer Reisenzein reviews empirical replication attempts and nearby inquiries over the two preceding decades: he declares that evidence holds up only for the first hypothesis, that "misattributed arousal from an extraneous source intensifies emotional reactions" (Reisenzein, 1983, 239). As Paul Griffiths points out, one would expect subjects to confabulate explanations of their abnormal, epinephrine-induced physiological arousal; therefore, Schachter and Singer's results "do not discriminate between this null hypothesis and the hypothesis that subjects were observing the normal arousal associated with the emotions they reported" (Griffiths, 1997, 82-83). In fact, proponents of Jamesian theories would predict this, since the Jamesian *specificity* claim does not disallow shared patterns of arousal between phenomenally distinct emotional episodes. That specific patterns of bodily activity trigger phenomenologically specific experiences does *not* imply that such states cluster into discrete affective kinds. James himself argued against rigidly taxonomizing discrete emotions based on overt physiological activity, pressing that such endeavors are futile—one emotion shades into another—and, therefore, that any divisions made between *kinds* of emotions must be arbitrary.²³

The *specificity* claim is therefore left unharmed by the experiment's findings. Although we can expect situations in which unspecific bodily arousal will be misattributed to a self-reported emotion, nothing in Schachter and Singer's findings suggests that different emotional experiences are *not* specified by different emotional reactions in normal circumstances (i.e., emotion episodes that aren't artificially induced in a laboratory setting). The *specificity* claim is revisited in Chapter 4, wherein evidence and arguments in support of it

²³James (1890, II: 449) writes: "The trouble with emotions in psychology is that they are regarded too much as absolutely individual things". Phoebe Ellsworth (2014) discusses this point with respect to contemporary treatments of James. I also raised the issue in Chapter 1, section 1.2.1., and defend a similar claim in Chapter 5.

are presented.

The experiment's design has also come under scrutiny. Martha Nussbaum (2001, 98fn) states that the authors confuse emotional contagion between stooge and subject with shared emotional responses. She also criticizes the execution of the experiment insofar as it is meant to capture states of euphoria via ad hoc play time in a room with a stranger. (Should we expect different emotional responses between introverted and extroverted subjects?) On top of all that, we can debate the merits of generalizing toward a universal theory from data on exclusively male undergraduate subjects till the cows come home.

Overall, then, we have numerous reasons to doubt the explanatory strength of the experiment's findings. Indeed, in their analysis, Julien Deonna and Fabrice Teroni say that "few now believe that [Schachter and Singer's] results can be taken at face value" (Deonna & Teroni, 2017, 57).

We can also question the extent to which Schachter and Singer's experiment engages with James's theory (and, by extension, the primitivist and other neo-Jamesian theories). Joseph LeDoux marks out the distinction between the explananda of Jamesian/neo-Jamesian projects and Schachter and Singer's project: "[Schachter and Singer] tried to explain how we deal with emotional responses once they occur... but did not give an account of what generates the responses in the first place. ... The brain's emotional business is... well underway by the time Schachter and Singer's mechanism kicks in. So what happens first" (LeDoux, 1996, 49)? The job of a Jamesian theory of emotion is to answer that more fundamental question.

2.2.3 Issues with Schachter and Singer's theory

Regardless of the numerous reasons for skepticism regarding the interpretation of the experiment's findings, we should not prematurely give up on Schachter and Singer's cognition-arousal theory itself unless we find convincing counterarguments or counterevidence. In fact their theory still holds sway over much of the field of emotion research despite widespread

acknowledgement of the experiment's shortcomings.²⁴ After all, just as a successful experiment does not definitively prove a theory, an unsuccessful experiment does not disprove a theory. While I cannot provide definitive evidence to disprove the theory's major claims, I do propose that we have good reason to believe that it does not adequately capture the essence of *emotion*. The problem, as I see it, is that the cognition-arousal theory is at once too cognitively demanding, yet too vague, to enable us to track and describe emotion.

Let's start with the issue of vagueness. Emotions, on the Schachter-Singer view, are constituted largely by emotion concepts. At least some, if not all, of these concepts will vary between cultures, languages, and generations. While the Japanese can sincerely and genuinely report experiences of *amae*, non-Japanese speakers cannot (Doi, 1971). Some words describing emotions or phenomena closely related to them are relative newcomers to English: *empathy* was introduced by psychologist E. B. Titchener as late as 1909; *schadenfreude* started appearing in English texts a few decades prior. And concepts are arguably better grasped by an individual as that individual develops and gains knowledge of those concepts. Perhaps, then, the people most capable of experiencing emotions (*qua* self-reports) are those who graduate with degrees in literature and foreign languages. (But then, I majored in literature and Japanese, yet here I am trying to work out what emotion is.) The point here is that if our emotions depend upon our understandings of emotion concepts, and if we only vaguely understand some emotion concepts, then we can only vaguely experience the emotions that depend on those concepts.

Furthermore, if emotion concepts vary across cultures, languages, and generations, then any attempt to understand biological or physiological contributions to emotion will be pushed in one of two directions: it will be restricted to explaining the role of bodily arousal (which Schachter and Singer downplay), or it will be forced into a framework of either evolutionary psychology, conceptual nativism, or embodied cognition so as to ground explana-

²⁴See for example the recent special edition of *Emotion Review* (volume 9, number 1) devoted to the cognition-arousal theory.

tions of (some) emotion concepts in terms of pan-cultural innateness.²⁵ As yet, none of these approaches has produced a compelling account of concept acquisition or generation that can fully explicate a Schachter-Singer-type cognition-arousal theory of emotion. For those who already accept that theory, they also must concede that emotion research will play second fiddle to research into the nature of concepts. This situation renders the cognition-arousal theory insufficiently detailed at present to provide a thorough explanation of emotion.

Now let's look at the demands. Schachter and Singer's theory relies on the emoter's ability to metacognize: to reflect upon one's own mental states, both affective (e.g., feelings of physiological arousal) and conceptual (e.g., beliefs), and subsequently label one's current emotion episode according to one's stock of emotion concepts (however they're acquired): *anger, joy, sadness*, and so on. Again, it's the event of labelling (i.e., generating a self-report of emotion state) that Schachter and Singer consider an emotion proper. Putting to one side the issues surrounding emotion concepts, two headaches remain with this conception: the *problem of metacognition* and the *problem of the concept of self*.

2.2.3.1 The problem of metacognition

Firstly, the cognition-arousal theory apparently denies emotions to any creatures that lack the capacity or resources to *metacognize*: to analyze and label their mental states. Pre-linguistic infants and almost all non-human animals are likely to be cast out of the realm of emotionality. This is a general issue for cognitivist theories, as I explained in the previous chapter; the cognition-arousal theory is one version of cognitivism. It's also a thoroughly undesirable consequence for any respectable theory of emotion. Indeed, the idea that beasts and babies don't experience emotions "has long since been reduced to an historical curios-

²⁵For a sample of perspectives from evolutionary psychology, see Al-Shawaf et al. (2016), Panksepp (1998), and Tooby & Cosmides (2008). The most notorious proponent of conceptual nativism is undoubtedly Jerry Fodor (1975); for discussions see Carey (2009) and Laurence & Margolis (2002). Recent attempts to ground emotion concepts in embodied cognition theories include Barsalou (1999), Niedenthal (2008), Niedenthal et al. (2014), and Prinz (2002, 2004). The notion of embodied emotion concepts is explored in Chapter 4, subsection 4.3.1.

ity” (Deigh, 2010, 27). Those of us who plainly see joy expressed in a dog’s pricked ears and wagging tail know that a theory of emotion must account for that dog’s experiences alongside our own. Let’s call this the *problem of metacognition*.

This is much less of a problem if we distinguish between emotions proper and self-reports thereof. If we grant that emotions logically and causally precede emotion self-reports, then we can accept that non-linguistic creatures may experience emotions even if they can’t publicly acknowledge that they do. When my dog wags his tail at me around dinner time or when I motion towards his leash, he experiences joy in anticipation. (On a Jamesian account, his feelings of his tail wagging partly causes his emotional experience of joy.) But he probably isn’t deliberately *telling* me so by wagging his tail. It’s not the equivalent of a verbal report of his joy, since presumably he lacks the capacity to metacognize.

The problem of metacognition disappears if we restrict the scope of the cognition-arousal theory: it does not account for emotions, but only reports thereof. It’s absolutely fine to say that one must metacognize to provide a report of one’s emotion state. (Indeed, that will be the focus of Chapter 4). However, conflating emotion with reports thereof renders the former too cognitively demanding, and too theoretically loose, to account for cross-species similarities. Claiming that metacognition constitutes emotion sees the tail wag the dog.

The Jamesian approach, in contrast, does not face the problem of metacognition since it does not require any reflection upon one’s emotional experiences to ground emotions. As James argued, emotions *just are* the feelings of bodily changes—albeit those changes that come about under certain situations reflecting organism–environment relationships. The next chapter will offer a theory of how such emotions come about.

2.2.3.2 The problem of the concept of self

The second, and stronger, reason for skepticism toward the Schachter-Singer theory (*qua* theory of *emotion*) is that apparently only creatures possessing a concept of self can qualify

as emoters. Consider what must happen in order to generate an emotion on this picture. An organism perceives (or imagines, or recalls) its situation. It forms a cognitive judgment (belief) about its situation with respect to its well-being—a belief that its life is at risk, or that it's about to receive some benefit, or that it has just lost something important. It also feels a general, vague sensation of arousal in its body. Noticing both of these mental events at the same time (or near enough to one another), it forms a second cognition: a belief that (i) both the evaluative judgment and the sensation of arousal are attributable to the external cue (i.e., the emotion stimulus), and (ii) identifying one's self as the subject of both the evaluation and the arousal (see also Gordon, 1978). Encountering the bear in the woods, I will judge that it is a threat to my well-being and feel a shock of arousal throughout my body. When I attend to both mental events, I link them together and label my overall state as an instance of fear. If this is a general picture of emotion, then access to a concept of self—one that can be employed as the subject of a belief—is a prerequisite for a capacity for emotion. These demands for rich cognitive capabilities, including an explicit concept of self, force further restrictions on the range of creatures that can experience emotions. Again, this looks grim for beasts and babies (unless, perhaps, they can pass the mirror test). If such creatures cannot attribute their mental states to themselves, then they cannot be said to experience emotions. Let's call this the *problem of the concept of self*.

But perhaps the problem of the concept of self is artificial, an exaggeration of the Schachter-Singer account. Indeed, the clearest evidence suggesting that the authors take a concept of self to be necessary for emotions is found in their experimental design. Data derives from subjects' answers to crucial questions: "How irritated, angry or annoyed would you say *you* feel at present?" and "How good or happy would you say *you* feel at present?" (Schachter & Singer, 1962, 387, my emphasis). One could argue that this limitation of the experiment should not similarly limit the theory it seeks to support. Maybe the theory could allow for emotions grounded in metacognition but lacking an explicit concept of self. Peter

Carruthers (2009; 2011) has developed an empirically informed theory that sees metacognition dependent upon mindreading capabilities. Briefly, metacognition results from reading one's own mind. Certain dissociative conditions may lead to an inability to identify one's concept of self while mindreading capabilities remain intact. One could, then, perform mindreading operations upon one's own mind without attributing the outcomes to an internal agent. Imagine an embodied organism whose cognitive architecture includes states such as beliefs, desires, and affects, but that lacks access to a concept of self, i.e., a non-external agent to whom its cognitive, sensory, and affective states can be attributed. Such a creature could, on this loosened cognition-arousal theory, label a cognition-sensation pair as an emotion without thereby claiming that it is experiencing an emotion: not subjectively "I'm sad", but objectively "That's sad", or even "That's a sad state of affairs". In English we reserve a small list of emotion concepts for such use in describing situations rather than agents. In Japanese the distinction between agent- and situation-describing emotion terms is even fuzzier: the word *kanashii* (sad), like most other emotion terms, is more commonly heard as a lone utterance than within a sentence predicated to oneself. The degree to which such language is metaphorical is up for debate. That humans can and do refer to an explicit concept of self in their emotion reports does not entail that every instance of emotion must involve a concept of self.²⁶

2.2.4 How to feel yourself

On the other hand, there is a strong sense in which all emotions—even the most primitive—*do* involve the self. It's one thing to have a concept of the self, but it's another to *feel* one's self. The notion of a pre-cognitive *feeling-self* is present in James's works: as Chapter 1 highlighted, James's psychology is intrinsically personalist. At the core of all mental states, for James, is a "nuclear self" made manifest by the permutations of the body at any given

²⁶None of this is to deny that emotion concepts aren't crucial to self-reports of emotions. Again, self-reports of emotions are not themselves emotions. A positive analysis of emotion concepts is offered in Chapter 4.

moment.

Similar notions of an implicit self also pop up in various contemporary philosophical and psychological theories. For instance, in writing about mindreading, Marc Jeannerod sits the concept of self (*narrative self*) and the feeling-self (*embodied self*) at different levels of consideration:

One of these levels is that of the narrative self. As a narrator, we obviously know who we are, where we are, what we are presently doing, and what we were doing before. Unless we become demented or amnesic, we have a strong feeling of continuity in our conscious experience. We rely on declarative memory systems where souvenirs (albeit distorted) can be retrieved and used as material for verbalization or imagination. (Jeannerod, 2005, 148)

Variations of the concept of self as narrator can be found throughout philosophy and psychology. I won't dwell on it here, and will instead attend to Jeannerod's notion of *embodied self*. He continues:

At variance with the narrative self, the type of self-consciousness that is linked to the experience of the embodied self is discontinuous: it operates on a moment-to-moment basis and is bound to particular bodily events, like actions. ... In other words, the embodied self mostly carries an implicit mode of self-consciousness, whereby self-consciousness is around but becomes manifest only when required by the situation. The related information has a short life span and usually does not survive the bodily event for very long. (Jeannerod, 2005, 148)

John Perry articulates a similar idea and attributes to the feeling-self an epistemic function that he calls *primitive self-knowledge*:

Primitive self-knowledge is what we get when Nature harnesses information; that is, when the processes of evolution result in a system that (a) has a repertoire of actions that can promote a natural value, such as survival or reproduction, and (b) has the capacity to pick up information about the circumstances of success of this action, and (c) has an architecture that harnesses that information. (Perry, 2017, 9)

Perry nicely illustrates primitive self-knowledge by describing a chicken looking at a piece of food. That the food is nourishing to the chicken, that such information is picked up by the chicken's visual system, and that the chicken is hungry (apparently they almost always are, Perry says), sees the chicken in a state that leads to its digesting the food. Thus, the "architecture of chickens harnesses information for the end of survival of the chicken in question" (Perry, 2017, 10). Enter primitive self-knowledge:

The chicken has what I call primitive self-knowledge. The information that the chicken detects is information about *itself*—using self in a perfectly modest way, and not as a hidden metaphor for selves or souls or complex minds of the sorts that humans have that allow them to admire themselves, hate themselves, and have identity crises. It is information about the the direction and distance of the kernel of corn from the very chicken who sees it that is harnessed. And it is harnessed for the benefit of that very chicken—it is the one which gets nourished as a result of pecking in that situation. So the chicken does have information about itself. It has *primitive* self-knowledge. This does not require mastery of the first-person, or some special inner version of 'I.' It just requires an architecture that allows pick up of information about the chicken, and harnesses that information to cause actions by that chicken, that promote the relevant values. (Perry, 2017, 10, original emphasis)

As we will see in the next chapter, primitivist emotions fit this description perfectly. Trig-

gered during information pick-up (via sensory registration), an emotional experience functions to motivate the organism to respond to its environment in a way proportional to its internal status. It is precisely by way of an emotional experience that creatures—chickens, dogs, humans—are able to feel themselves (or feel their selves). Primitivist emotions yield primitive self-knowledge.

Coming back to our fictional self-less emoter (above on page 77), we can now see that it would inevitably experience a feeling of self woven through its emotional experience. Given the capacity to metacognize, it would analyze its own phenomenal and cognitive states and produce a sincere statement of what it finds therein. However, lacking an explicit concept of self, it would still produce an objective, externalized emotional report: “That’s a sad state of affairs” or similar. Such an utterance would fail to convey the feeling-self quality of the event it represents. Its intentionality would be misdirected if not missing. That quirk does nothing to deny the involvement of a *feeling-self* in the underlying emotion, however. The self merely gets lost in translation.

This gives us further reason to divorce emotion from self-reports thereof. Our self-reports can be limited in unique ways while our phenomenal experiences are unrestrained. Cognitively unsophisticated creatures—beasts and babies—can still feel emotions, and those emotions can carry information about their selves as embodied agents, despite their lacking explicit concepts of self.

The problems faced by Schachter and Singer’s cognition-arousal theory of emotion are demonstrative of the general tension between Jamesian and cognitivist projects. James’s theory is fundamentally noncognitivist, and that is one of its major strengths. It allows for emotions to exist as simple phenomena. And because it asserts that emotional experiences do not require a stock of emotion concepts, James’s theory enables creatures of drastically different psychological capacities to experience them. But by insisting that the term *emotion* tracks something neatly articulated by linguistically capable humans, Schachter and Singer

deny emotion's status as a fundamental element of a psychological ontology.

Schachter and Singer's cognition-arousal theory, then, is psychology's equivalent of a mule: a hybrid of two disparate theories that manages to pull some weight but ultimately cannot reproduce.

2.3 What is a neo-Jamesian theory of emotion?

As I mentioned earlier, my theory derives from James's claim that an emotion is the feeling of bodily changes as they occur in response to some present stimulus (see the previous chapter). I also mentioned, though, that I'm not the first person to propose a neo-Jamesian theory of emotion. The next two subsections (2.3.1–2.3.2) discuss two recent attempts to revive the Jamesian approach. Section 2.4 shows where they both fall short of a faithful reconstruction of James's theory.

2.3.1 Antonio Damasio

In 1994 neuroscientist Antonio Damasio published his book *Descartes' Error*. This book has become immensely popular even among nonacademic readers. (I bought my copy at an airport bookstore. It was sitting on the "bestsellers" shelf.) It has been hugely influential within academia as well: according to a cursory Google Scholar search, to date (August, 2018) it has been cited just shy of 27,000 times. Around the half-way point of the book Damasio introduces a couple of William James's claims from his chapter on emotions in the *Principles*, and cites them as describing "the mechanism essential to understanding emotion and feeling" (Damasio, 1994, 129). In doing so, Damasio sets himself up to be regarded as a neo-Jamesian regarding emotions.

Though he sees in James the fundamental pathway to explaining emotions, he nonetheless takes issue with aspects of James's theory as he interprets it. The first issue is that James

“gave little or no weight to the process of evaluating mentally the situation that causes the emotion” (Damasio, 1994, 129-130). His implication is that, though James’s theory can explain some of the emotions we experience in early life (and, presumably, many of the ranges of emotions experienced by other animals), it cannot explain the kinds of emotions triggered by the contemplative and convoluted situations that the typically developed adult undergoes. (This echoes the complaints of the cognitivists described in section 2.1.1.)

The second issue Damasio highlights is that James did not leave room for anything other than actualized, extraneural bodily changes to stand as causes of the felt quality of an emotion. In our terminology, Damasio disagrees with James’s claim that an emotional experience is necessarily dependent upon a whole-body emotional reaction. The combination of these issues, on Damasio’s reading, leaves James’s theory insufficiently nuanced to account for the range of phenomena that demand explanation. Damasio presents James’s theory in a simple characterization:

In short, James postulated a basic mechanism in which particular stimuli in the environment excite, by means of an innately set and inflexible mechanism, a specific pattern of body reaction. There was no need to evaluate the significance of the stimuli in order for the reaction to occur. Matters were not made more clear by his lapidary statement: “Every object that excites an instinct excites an emotion as well.” (Damasio, 1994, 130)

Damasio’s interpretation of James is hardly immune to scrutiny. Indeed, the picture painted in the previous chapter is a far cry from an inflexible reflex mechanism: Jamesian emotional reactions are trainable and malleable (see also Hatfield, 2007, 421; Barbalet, 1999). And in the same year that Damasio’s words here saw print, Ellsworth published her paper in which she presents James as an appraisal theorist (Ellsworth, 1994; discussed in the previous chapter). While Ellsworth pushes James too far towards cognitivism, Damasio, conversely, is quick to pigeonhole James’s noncognitivist stance on emotions into a rigid, lever-and-

pulley-like mechanism.

Despite his misinterpretations of these key features of James's theory, Damasio builds his own theory on the most general foundations James had laid down a century prior. He splits the range of phenomena called *emotions* into two kinds: primary emotions and secondary emotions. Primary emotions, he explains, should be thought of in the sense that James (per his reading) describes them (Damasio, 1994, 131-132): many animals, including humans, may be wired to respond to certain features of stimuli that they perceive in the world, such as size, span, type of motion, ranges of sounds, and certain bodily states such as physical pain. The brain's limbic system (the commonly postulated but variously demarcated system responsible for producing emotions) would process these perceptual signals and trigger a pattern of bodily changes constitutive of an emotional reaction (e.g., a fear reaction) which would subsequently be felt as an emotion.

Of the perceived stimuli, Damasio writes:

Note that in order to cause a body response, one does not even need to "recognize" the bear, or snake, or eagle, as such, or to know what, precisely, is causing pain. All that is required is that the early sensory cortices detect and categorize the key feature of a given entity (e.g., animal, object), and that structures such as the amygdala receive signals concerning their *conjunctive* presence. (Damasio, 1994, 131-132; original emphasis)

As far as consciousness is concerned, Damasio states that one feels one's emotion "in connection to the object that excited it", such that an emotional experience is "the realization of the nexus between object and emotional body state" (Damasio, 1994, 132). (This is not far from the more radical position Ellsworth (1994) attributes to James.) And the function of an emotional experience is to buy the organism time: as attention is diverted to the cause of the emotion, the conscious organism is better placed to control its subsequent behaviors than it would be were it merely to respond in a rigid, reflexive manner. Emotional experiences

can also lead us to associate stimuli in such a way that we can extend our dispositions to react with an emotion to repeated exposures to a stimulus kind (Damasio, 1994, 132): for instance, one encounter with a snake resulting in a fear reaction may be sufficient to dictate fear reactions upon subsequent encounters with snakes.

Damasio presents the above as a plausible explication of what he calls a “Jamesian emotion”, a kind of emotion that is innate and preorganized (Damasio, 1994, 132). He offers this mechanism to account for primary emotions—the range of emotional phenomena that do not have cognitively or socially complicated antecedent causes—but goes on to argue that secondary emotions, namely those unique to cognitively sophisticated creatures such as adult humans, require a different set of mechanisms. Secondary emotions emerge through the following sequence (Damasio, 1994, 136-138). To begin the episode, the organism consciously deliberates over its state of affairs. Damasio explains this in terms of a sequence of mental images, grounded in multiple sensory modalities, that constitute a “cognitive evaluation” of the situation. The evaluation then triggers a representation of how such a situation has been responded to previously in the organism’s individual history. That is, the evaluation is associated with prior emotional responses to similar situations, and these associations are ultimately informed by episodes of primary emotions. Simultaneously, certain areas of the limbic system trigger the kinds of bodily changes stereotypical of corresponding primary emotional reactions (such as visceral changes and release or inhibition of hormones and neurotransmitters).

Thus, the body makes its way into the picture as producer of the feeling of a secondary emotion. However, it is primarily the changes in levels of various neurotransmitters that directly accounts for the feeling of an emotion. Of course, the extraneural body contributes plenty of important neural signals to the brain that manipulate the quality of the feeling, and it also manipulates the way in which neural signals are communicated on a slower timescale, when the hormonal glands release their chemicals into the bloodstream and those chemi-

als penetrate the blood-brain barrier. In fact, Damasio takes the brain to be primary in monitoring what is going on inside both the brain and the rest of the body, along with monitoring what is going on in the world (Damasio, 1994, 90). A feeling is the manifestation in consciousness of the state of the body:

As body changes take place, you get to know about their existence and you can monitor their continuous evolution. You perceive changes in your body state and follow their unfolding over seconds and minutes. That process of continuous monitoring, that experience of what your body is doing *while* thoughts about specific contents roll by, is the essence of what I call a feeling. (Damasio, 1994, 145; original emphasis)

As for specifically *emotional* feelings, he continues:

If an emotion is a collection of changes in body state connected to particular mental images that have activated a specific brain system, *the essence of feeling an emotion is the experience of such changes in juxtaposition to the mental images that initiated the cycle*. In other words, a feeling depends on the juxtaposition of an image of the body proper to an image of something else... (Damasio, 1994, 145; original emphasis)

Here we see an interesting feature of Damasio's theory: intentionality. Given that emotional feelings result from a relationship between the body (or state thereof) and the stimulus objects mentally portrayed, Damasio takes it that such feelings are *about* their stimulus objects (Damasio, 1994, 147-148). At the same time, feelings are informative of the state of the organism's body, and thereby carry a secondary intentional character. In other words, the feelings of "Damasioan" emotions are at once about both stimulus and body.

Damasioan emotional feelings therefore satisfy the first two of the three desiderata that we might fairly expect of a proposed theory. Do they also satisfy the third criterion: moti-

vation? Perhaps. Recall that the function of Damasioan emotional feelings is to allow the organism to take control of its behavior with respect to the stimulus situation. If we are compelled to understand motivation in the sense that the organism is prompted to consider its options, then yes, emotional feelings are motivating. However, if we understand motivation as prompting world-directed behavior, then it does not appear that emotional feelings are motivational since they are disruptive, not irruptive, with respect to our innately determined behavioral responses to stimuli. Conscious feelings buy us time to reconsider our options rather than blindly act according to our instincts. One might consider this their *demotivating* quality. On the other hand, the fact that feelings prompt us to think means that they motivate an internal change of mental state; even if this is not an externally expressed behavior, technically it results from feelings' motivational aspect.

Finally, note that Damasio disagrees with James on what we should identify as the emotion proper. While for James an emotion is nothing more than the *feeling* of bodily changes, Damasio identifies an emotion as "a collection of changes in a body state". What James calls an *emotional experience* is approximately akin to what Damasio calls an (intentionally characterized) *emotional feeling*. I'll return to Damasio's account of feelings again shortly, but this is enough detail to demonstrate that Damasio's position is an outgrowth of the James-Lange theory, and not James's theory proper.²⁷ Now we can turn to another, similarly inspired theory of emotion; one that finds its roots planted not only in the James-Lange postulate, but also in theories contrary to James's own position.

²⁷The previous chapter detailed numerous distinctions between James's and Lange's theories and argued that we should dispel the James-Lange postulate from our interpretations of James's corpus. Lange took the physiology to be primary in an explanation of emotional phenomena. This feature carried over to the James-Lange conflation. James, however, explicitly denied this: again, he identified the *feeling* as the emotion proper. Thus, since it emphasises neurophysiology for explanatory purposes, Damasio's account is much closer to the James-Lange account, or even just Lange's theory.

2.3.2 Jesse Prinz

A decade after Damasio's book appeared, Prinz published his own "neo-Jamesian" emotion theory in a book titled *Gut Reactions* (2004). Prinz explicitly cites "the writings of James and Lange" as ancestors of his theory (Prinz, 2004, 224). But his theory offers a substantial overhaul of its source material, bringing in a precisely defined cognitive element of appraisal. Though Prinz rejects Ellsworth's appraisalist interpretation of James (Prinz, 2004, 5fn), he also rejects the idea that appraisals need to be thought of in terms of propositional thoughts. In attempting to find a theoretical middle ground between the James-Lange, body-first position and the cognition-dependent appraisalist school, Prinz defines emotions as "embodied appraisals".

He derives his notion of appraisal from theories of emotions that developed as attempts to comfortably sit emotions either alongside or within cognition. I gave a description of some of appraisalism's major features in the previous chapter. The particular version of appraisalism that Prinz relies upon is due to Richard Lazarus, a psychologist who developed a cognitivist theory of emotions between the 1960s and 1990s. The mature theory Lazarus presents in his book *Emotion and Adaptation* (1991) is grown from ideas he presented a decade prior in a debate with fellow cognitive psychologist Robert Zajonc. Zajonc was the first to strike, claiming that preferences are immediately present to an organism by way of affect; in fact the subtitle of the paper that catalyzed the debate reads "Preferences Need No Inferences". If an organism's preferences drive its emotions, and preferences do not require any inferences or "mental work" (read: cognitive processing) in order to come about, then emotions need not rely on any cognitive causes (Zajonc, 1980).

Lazarus's response sees emotions as underpinned by cognitive processing (Lazarus, 1982), and by 1991 he has articulated his theory of emotional appraisal. As Prinz reads him, Lazarus "defines appraisals as evaluations of what one's relationship to the environment implies for one's well-being" (Prinz, 2004, 14; cf. 52). Lazarus identifies six dimensions at work

in appraisals which can be sectioned into two groups: three dimensions whose function is to “establish that something is emotionally significant” (Prinz, 2004, 14), and three that “pertain to the resources one has available for coping” (Prinz, 2004, 14). Each of these is an evaluative judgment which, when collectively summarized, produce a *core relational theme* (Prinz, 2004, 15). Core relational themes are representations of organism-environment relations. For any given emotion, positive or negative, one can find (in theory) underlying it a specific core relational theme (Lazarus, 1991, 121). The core relational theme underlying anger, for instance, Lazarus describes as “a demeaning offense against me and mine”; that underlying guilt is “having transgressed a moral imperative”; and so on (see Lazarus, 1991, 122, table 3.4).

As Lazarus’s view stands, at least via Prinz’s reading of it, a core relational theme is at the heart of the process underlying both an emotional reaction and an emotional experience. In building his novel theory, Prinz takes Lazarus’s notion of appraisal as producer of core relational themes but attempts to strip it of its cognitivist machinery. The inner mechanisms of the Prinzi appraisal mechanism do not operate over cognitive judgments of perceptual content (*contra* Lazarus’s theory) but are themselves perceptual processes. More specifically, Prinz replaces the propositions at work in cognitive appraisal with mental representations that function to detect bodily changes (Prinz, 2004, 52). In registering changes in the body, emotions represent core relational themes.

Here is an example he provides for how the sequence of an emotion plays out:

Consider the chain of events leading to fear. Something dangerous occurs. That thing is perceived by the mind. This perception triggers a constellation of bodily changes. These changes are registered by a further state: a bodily perception. The bodily perception is directly caused by bodily changes, but it is indirectly caused by the danger that started the whole chain of events. It carries information about danger by responding to changes in the body. That further state is

fear. (Prinz, 2004, 69)

One can compare this sequence with the standard interpretation of James's view and find fairly little in way of disagreement. But Prinz offers an elaboration on the standard Jamesian (or James-Langean) picture when he claims that the experience of fear "carries information" about the danger to which the body is reacting. On the Lazarus-style, cognitive appraisal view, that information (the core relational theme) would come in propositional form, as an evaluative judgment. The crucial departure from cognitivism here is that the information is carried not by propositional judgments but by *feelings*: "Feelings can obviate the need for cognition, because feelings carry information. The discrete motions of our bodies convey how we are faring in the world" (Prinz, 2004, 78). In other words, the phenomenality of an emotional experience does the work of representing the particular core relational theme with which it co-occurs.

Prinz maintains that core relational themes can be represented in the mind without having a propositional structure (Prinz, 2004, 65). Part of what is represented is something external to the organism—the thing that is in some represented way related to it—and that something need not be represented *as* what it is, per se: "By analogy one might say that a state in the visual system *registers* a particular luminance discontinuity, but it *represents* an edge" (Prinz, 2004, 58; original emphasis). In the same way, a registration of a specific bodily change represents a core relational theme of a specific kind—even if the nature of that core relational theme remains opaque to introspection: "We can form the judgment that there has been an irrevocable loss, but we seldom do. Sadness can occur without that judgment. But sadness represents what that judgment represents. It has the same meaning but a different form" (Prinz, 2004, 65). Prinziian emotions, then, are mental representations, caused by bodily changes, that function to represent a core relational theme depicting the present organism-environment relation. The bodily changes themselves follow from the perception of a core relational theme—the perception of the relation the organism finds itself in with

respect to its surroundings.

Thus, Prinz arrives at his claim that emotions are “embodied appraisals”. They are appraisals in that they represent situations that concern the organism (its core relational themes), and they are embodied in that they are caused by changes throughout the body.

But there is an issue to address in Prinz’s theory. Consider that block quotation from Prinz, on page 89, which lists the steps underlying the experience of fear. “Something dangerous occurs”, he writes, and the mind perceives *that thing*. There’s a gap in the explanation here; one that requires quite a lot of work to fill. Precisely what is perceived? And what is it about that thing that triggers the bodily changes that inform a fear response? Prinz notices this gap himself: “A slithering snake cannot cause one’s skin to crawl without a mediating link in between. There must be some inner state that detects the snake and then causes the physiological change to take place. Emotions must have inner causes” (Prinz, 2004, 74). He doesn’t offer an argument for why there *must* be an inner state that triggers an emotional response; he merely takes the assumption for granted. But as I showed in the previous chapter, a legitimately Jamesian position can make do without invoking inner (mental) states in order to bring about physiological change. Nonetheless, let’s grant him his claim and put aside the issue of mental causes for now (but we’ll come back to it shortly). He continues, postulating that snakes trigger fear in humans when we perceive them and activate “a primitive visual representation of the snake”. This representation then mediates “between the external danger and the racing heart” (Prinz, 2004, 74). The inner cause of one’s fear following the perception of a snake is a perceptual representation of that snake.

While there is plenty of evidence that snake phobias are easily instilled or triggered in primates, we cannot account for all of our emotional responses in this way. Prinz attempts to alleviate this concern by offering associative learning as the mechanism by which subsequent interactions with situations warranting emotional responses come to invoke those emotions. An instance of consuming noxious food (e.g., spoiled milk) can lead to a disgust response to

that kind of food even if it's not noxious (e.g., fresh milk). "Memory forges links between emotions and representations of the particular objects that elicited them" (Prinz, 2004, 75). This appeal to association is similar to that seen in Damasio's theory above (page 85), and in James's own (see the previous chapter).

Notice Prinz's placement of an object at the center of emotion-inducing perceptual content. Fear of snakes comes from the detection of the snake, which triggers a mental representation of *the snake*. It's something about the object type (snakes, in this case) that triggers the emotion (fear, in this case). Yet Prinz does not offer us any story of the way in which perception leads to bodily changes—especially the sorts of adaptively beneficial, complex, and coordinated bodily changes we could expect of our emotional responses! Without filling in those early details of an emotion's causal story, we can't arrive at a complete explanation.

To be fair, Prinz does in fact offer the seeds of an explanation later in his book. In defending emotions' status as perceptual in nature, he reminds his readers that emotions are perceptions of bodily changes, and further defines those bodily changes as "the body's preparation for action" that "enable us to behave in appropriate ways". "In this sense," Prinz continues, "one might think of emotions in Gibsonian terms" (Prinz, 2004, 228). Here Prinz is adopting James J. Gibson's famous offering of affordances as the contents of perception (Gibson (1979)). He gives us a brief, theoretically shallow description of affordances: "Gibson... says that in ordinary perception we perceive the actions afforded by the objects in our surround. We see that a chair affords sitting and a hammer affords wielding. Emotions are perceptions of affordances in this sense. By registering bodily changes, emotions allow us to literally perceive that situations afford a range of possible behavioral responses" (Prinz, 2004, 228).

It seems that Prinz is here equating Gibsonian affordances with the core relational themes he adopts from Lazarus. I think that Prinz is correct to invoke Gibson, and in fact I will be doing the same myself in Chapter 3. However, the ways in which Prinz and I arrive at our

Gibsonian positions differ markedly. Prinz is here appealing to Gibson out of theoretical convenience: talk of affordances provides a nice heuristic for him to make his point about the motivational feature of emotions. That is, Prinzian emotions motivate us towards the behaviors of which they inform us we are capable. My own position, meanwhile, logically depends on Gibsonian ecological psychology in a stronger sense. In fact, I'll go so far as to argue that my theory's dependence on Gibson's contributions is a conceptual outgrowth of its Jamesian roots.

Before moving on we should assess Prinz's theory against the three desiderata demanded of a theory of emotion. Prinzian embodied appraisals are informative, in that they represent how our bodies are faring in the world along with opportunities to act in response. They are intentional, in that they represent a relationship between the stimulus object and the body. And they are motivational, in the sense that they prompt us towards the possible reactions of which they inform us.

2.4 Do emotions require mental causes?

We have just seen that both Damasio's and Prinz's theories see emotional experiences as possessing intentionality: they are about relationships between the organism and features of its environment. However, the way this intentionality is secured, on either picture, relies on its being brought in via perceptual representations. Perceptual representations trigger changes across the body which are then represented in emotional feelings. So, while the body remains the primary cause of emotional experiences, perceptual representations, we can say, are the *occasional causes* of all emotions, at least on these accounts. Granted, Damasio's suggested list of world features that can trigger "primary" emotions is more elemental, so to speak, than is Prinz's: Damasio talks of object features such as size and span as stimuli, while Prinz explicates the necessity of perceiving an *object*, such as a snake, in order to trigger an emotion.

On the other hand, Damasio's triggering mechanism only works given its ability to utilize information stored in representations ("mental images") of prior encounters with objects displaying such features as he lists. Thus, Damasio's and Prinz's accounts both require mental causes for emotional responses. The explanatory sequence on either view jumps between descriptive levels: a psychological state causes physiological changes which cause another psychological state. A necessary conclusion of this construal is that emotional experiences are unique to creatures who have access to mental images or concepts. Put crudely: one must have a perceptual mind *before* one can feel anything like an emotion.

One can question whether this psychological precursor is necessary for all emotional reactions. The most commonly stated answer is yes: an affective state is not an emotion proper unless it has some object and is *about* that object. Sometimes this point is made to distinguish emotions from moods: where one has the symptoms (of fear, say) but cannot identify (through whatever measures are available) an object or situation that stands as the occasional cause of those symptoms, then one's affective state is a mood, not an emotion. James, however, made no such distinction, and in fact took objectless emotions to be evidence in favor of his claim that emotional reactions trigger emotional experiences. It matters not one bit to the classification of one's mental state whether one can find an occasional cause: if one feels the symptoms of fear, then one simply *has* the emotion of fear (James, 1890, II: 459). Since bodily states are the direct causes of emotions proper (*viz.* emotional experiences), distinguishing between emotions and moods based on the presence or absence of a psychological cause is both uninformative and arbitrary.

However, James does refer to objectless emotions as "pathological cases" (James, 1890, II: 458), and indeed in most instances one will identify one's emotional experience with reference to some occasional cause (the spider scurrying across your desk, or the unexpected presence of a loved one, or the like). So, there is nothing wrong, in principle, with claiming that a capacity for perception (or, perhaps, cognition) must logically precede a capacity for

emotion. But it's not the only option, even if it is most intuitively obvious one.

Another way to think of the relationship between emotionality and perceptual capacity is to reverse the order of constitution: the capacity to perceive the external world meaningfully emerges out of emotionality. I will argue for exactly this position in the next chapter. Briefly, the argument will take the following form:

- P1 A general physiological capacity for affect evolved before perceptual systems.
- P2 The development of mode-specific perceptual systems was guided by organisms' ecological concerns.
- P3 These ecological concerns were (and are) tokened as affective experiences: they are grounded in microscopic and macroscopic bodily changes (*emotional reactions*) on the part of the organism.
- P4 As physiological sensory systems developed greater modal specialities (e.g., sophisticated tactile, visual, and auditory pathways), correlative affective experiences helped to (i) inform and maintain a self-other divide and (ii) relate the ecological significance of external stimuli to the organism.
- P5 By enriching incoming sensory information with interoceptive information regarding the current state of the body (*emotional reactions*, e.g., hunger, thirst, vulnerability), a general affect system can produce a world-directed feeling: an *emotional experience*.
- C Thus, emotions need not depend upon prior perceptual or cognitive states to evoke them; they need only rely on interoceptive monitoring of exteroceptive activity.

This argument will form the basis of the *primitivist theory of emotion*. A key feature of the theory—one that derives from James's insights—is that emotion is a psychological prim-

itive: it is not reducible to other psychological events, including the physiological changes (*emotional reactions*) that trigger instances of emotion. Rather, emotion contributes to other psychological processes such as perception and cognition.

Conclusion

This chapter has examined extant responses to James's original theory of emotion. While the cognitivist approach, currently dominating the field of emotion research, is a conceptual counterposition to James's theory, some theorists have attempted to extend the spirit of Jamesianism into the contemporary sciences of the mind. Although these approaches have themselves furthered knowledge and guided empirical investigation into the nature and operations of emotion, they all rest on an assumption that emotions are necessarily triggered by some prior mental state: either a cognition or a perception. But this claim is not present in James's theory (see Chapter 1, section 1.2.2); and as the next chapter will argue in detail, nor is it logically necessary or probable according to our current understanding of the mind.

Indeed, James's theory *does* require alteration in order to bring it in line with our current knowledge. The message underlying the present chapter is that we must be careful to make the right kinds of changes if we are to protect and progress James's legacy. With that in mind, Chapter 3 will offer a positive, contemporary take on the Jamesian theory of emotion.

Chapter 3

A primitivist theory of emotion

Introduction

This chapter presents a novel theory of emotion. It is a *primitivist theory of emotion*. It presents emotions as affective phenomena that are intrinsic to the consciousness of creatures capable of sensing the world outside their bodies.

The primitivist theory of emotion draws on the lessons of the previous chapters while maintaining the spirit of James's theory: it is a neo-Jamesian theory of emotion. Like James's theory, primitivism (i) understands emotional experiences (kinds of subjective feelings) as emotions proper, (ii) identifies bodily changes as the direct causes of emotional experiences, and (iii) emphasizes the important causal and ecological relationship between emotions and other mental states such as perceptual and cognitive representations. At the same time, its major tenets are motivated by recent discoveries and theories from across the sciences of the mind.

To foreshadow discussion, the following causal schema, derived from James's theory (see Chapter 1, page 23), represents the neo-Jamesian account developed here:

STIMULUS → SENSORY REGISTRATION → BODILY EXCITEMENT (INTEROCEPTIVE

MONITORING)

→ EMOTIONAL EXPERIENCE

Note that this portrayal replaces the notion of stimulus *perception* with *sensory registration*. The difference is important. Sensory registration is not a mental process per se; it can be examined and described at the level of sensory organ physiology and neural activity. Perception, however, is a process that involves description of sensory registration at one level but also the generation of intentional mental states at another. This explanatory positioning of sensory registration is a crucial feature of my primitivist theory: it does not require that there be any prior perceptual processing in order for an emotion to be generated. Instead, as the present chapter will argue, this emotional sequence itself contributes to the generation of a perceptual state.²⁸

The chapter is comprised of two major sections. Section 3.1 explores affect: its experiential characteristics (3.1.1), bodily causes (3.1.2), and ecological functions (3.1.3). It then presents emotional experiences as species of affective states. Emotional experiences, which are emotions proper, are distinguished from other affective states according to their intentionality and the information they carry. While other affective states (such as hunger, thirst, and pains) portray the state of the organism's body, emotional experiences inform the organism of its possible behavioral responses to its environment.

In this respect, emotions contribute a certain affective quality to perceptual states: they represent certain kinds of *affordances* (Gibson, 1979). This idea is detailed in section 3.2. Many contemporary theorists advocate for a theory of perceptual affordances, but I believe the primitivist theory can benefit from revisiting its key concepts. This second section consequently differentiates between physical affordances, defined and described in the abstract, and psychological affordances, whose constitution requires the active involvement of the emotive, sensing organism.

²⁸I also want to remain silent for now on whether perception is direct or indirect (i.e., internally mediated). The current causal schema allows for this.

A caveat before we dive in. The primitivist theory demands a revisionist attitude toward the concept emotion. As the first and second chapters demonstrated, James's question, "What is an emotion?" lacks a clear and universal answer. Primitivism is not proposed as a resource for explaining all events that we humans dub "emotions". Moreover, it is a theory for *psychological* explanations of emotional phenomena, not universalist or folk explanations. Therefore, it may look a little narrow or weird in scope to those who have not spent much time working in the philosophy and the sciences of the mind. Some concerns that apply specifically to human emotions will be addressed in the next chapter.

A major issue is that we readily apply the term *emotion* and its cognates to so many phenomena that we struggle to identify anything that makes some phenomenon emotional by its nature. One response from a scientific perspective is to eliminate emotion from a scientific vocabulary (Griffiths, 1997, 2004a,b). But this is heavy-handed: according to a primitivist construal, emotion denotes mental phenomena that play important roles in explanations of human and non-human psychology. Therefore, proponents of the theory introduced below should revise, rather than eliminate, their use of the term.

Now let's get started on the theory.

3.1 Affect: its character, causes, and functions

A neo-Jamesian theory of emotion will adhere to James's claim that *emotional experiences*, the felt qualities of emotions, are directly caused by *emotional reactions*, the patterns of physiological change that occur during emotion episodes. We see this in Antonio Damasio's and Jesse Prinz's accounts in the previous chapter. It is also present (if in a more restricted way, or relying on different conceptual descriptions) in other theories such as those presented by Joseph LeDoux (1996), Giovanna Colombetti (2014), and Lisa Feldman Barrett (2017). The body enters the picture in different ways on all these theories: LeDoux, Barrett, and Dama-

sio, for instance, all emphasize that the brain receives inputs from the body and produces affective states that correlate with the interoceptive state informed by the body. (More on this below.) Prinz seems to hold a similarly neurocentric view, according to his recent book on consciousness (Prinz, 2012), though his story invokes more talk of representation than do the three scientists just mentioned. Meanwhile, Colombetti's enactivist theory sees the body play a much more direct role in the production of affective states; though what qualifies as affectivity, for Colombetti, is drastically different than what most affective scientists would readily accept (see Colombetti, 2014, ch. 1).

All of these theorists agree that emotional experiences cannot reduce to mere affective states: emotional experiences are something “deeper” or “richer” than affects, or perhaps they necessarily have objects whereas affects can be free-floating and intentionless. (This difference often grounds an explanatory distinction between emotions and moods, but see Prinz (2004, 182-188) for analysis.) But the differences between these theorists' accounts prompt queries of what is an affective state (such as an emotional experience), how it comes to be, and what are its functions. These three questions can be investigated more or less simultaneously. Once we have a clear understanding of the possible relationships between the body and affectivity, we can further explore the relationship between emotional reactions and emotional experiences.

3.1.1 The dimensions of affective qualities

The term *affect* is commonly employed in the cognitive sciences to label “raw feelings” or the experienced qualities of emotions, moods, and other such mental phenomena. Some theorists, especially Damasio, prefer to avoid using the term *affect* due to the vague role it plays in affective science: “The term *affect* is often used as a synonym of ‘mood’ or ‘emotion’ although it is more general and can designate the whole subject matter we are discussing here: emotions, moods, feelings” (Damasio, 1999, 341-342n10). Recall from Chapter 2 that

Damasio takes emotions to be patterns of physiological change. Therefore, he implies, the term *affect* is ambiguous regarding whether what it designates is a physiological or a psychological phenomenon. What *our* present interest is—emotional experiences—Damasio calls *feelings*.

Contrastingly, Barrett presents an account of affect that she understands as standard within affective science: affect is “simple feeling”, the “general sense of feeling that you experience throughout each day” (Barrett, 2017, 72). Barrett takes affect to be *the* fundamental component of the experiential side of mind. Either she eschews Damasio’s point about the ambiguous use of *affect* in affective science, or a conceptual consensus has been reached in the two decades since Damasio raised his concern. Going forward, I will use the term *affect* in Barrett’s narrower, strictly phenomenal sense. Thus, an emotional experience is an affective state that is brought about by an emotional reaction, which is *not* an affective state.

According to James Russell’s (1980) now-classic definition, all affective qualities can be described in terms of two dimensions: arousal and valence. Arousal is the state’s degree of excitation: how calm or how agitated you feel. Valence is the extent to which the affective state is pleasant or unpleasant. The affective state that comes from relaxing on a sofa after a long day might be described as *contentment* (low arousal, positive valence), while the state that follows the perception of a spider crawling up one’s leg might be described as *fear* (high arousal, negative valence). However, while many states are easily described as fitting into one of the four quadrants of this “circumplex model of affect” (Russell, 1980), others appear more ambiguous or difficult to place in a single quadrant. Is the experience of salivating at the sight of a not-quite-ready meal a positive or negative state? What about the feeling of muscular aches and fatigue following a strenuous gym session? Emotional feelings such as *bittersweet* and *nostalgia* are similarly difficult to place at any one point, and we often describe ourselves as holding “mixed feelings” towards situations. (Prinz (2004, 164-167) discusses some other examples.) Perhaps instead of thinking of valence as a single dimension

that can be either positive or negative, we should reconceive it as describing two possible dimensions (pleasure and displeasure). Arousal then would be the third dimension of the state space by which we can characterize a given affective state. Anyone who has endured the hilarious yet oppressive agony of a malicious tickler's attack may favor this proposed augmentation of Russell's model. (But see Russell, 2017 for a plausible counterargument.)

Another dimension that could be considered in descriptions of affective qualities is bodily localization. As we will see in Chapter 4 (section 4.2.2), when we report on our emotion states we often are able to locate (if vaguely) our feelings in different body regions in terms of increased or decreased activity. However, such acts of localization require goal-driven introspection and may not yield fully accurate results each time. Reporting on affective states (such as emotional experiences) is very different than unreflectively experiencing them. Attention will be directed differently under either circumstance. Therefore, it's difficult to determine whether bodily location factors pre-reflectively into the characterization of affect *qua* subjective experience. (Note that the body's phenomenological contributions and its generative contributions are distinct: while the body undoubtedly contributes to the generation of affect, it's not at all clear that any of an affective state's phenomenal character derives from embodiment or localizability.)

3.1.2 Bodily causes of affect

Orthogonal to the question of how to draw the dimensions of affectivity is the question of how bodies produce affective states. As we will see below, it is common to see affective states explained as the experiential effects of *interoception*, the detection of physiological changes. But filling in the details is hardly a simple undertaking: accounting for the physiological causes of mental events is just one iteration of the hard problem of consciousness (see Chalmers, 1995). Nevertheless, each of the authors mentioned above has attempted to do just that. Two recent accounts of the neurophysiological mechanisms underlying affect

are worth examining.

3.1.2.1 Damasio and Carvalho's neurological account of feelings

In a recent opinion article published in the journal *Nature Reviews Neuroscience*, Damasio and Gil B. Carvalho (2013) further the former's argument that feelings (*affects* in our terminology) are mental experiences caused by changes in body state.

The physiological changes that bring about feelings are *interoceptive* rather than *exteroceptive*. This means that feelings are not the direct result of, or psychologically or phenomenologically equivalent to, perceptual states derived from exteroceptive sensors (such as visual or auditory states or systems); the latter can "cause emotions and ensuing feelings but are not feelings in and of themselves" (Damasio & Carvalho, 2013, 143). (Moreover, the authors write, exteroceptors—eyes, ears, noses, and the like, with their corresponding neural structures—are likely to have evolved later than interoceptors (Damasio & Carvalho, 2013, 145).)

Damasio and Carvalho describe the causes of feelings in three ways: a macro-level structural description, a micro-level structural description, and an occasional description. At the macroscopic level they implicate numerous subcortical brain regions that are involved in the production of homeostasis-relevant changes. *Homeostasis* is an organism's baseline state of being in standard or ideal circumstances; it is the organism's bodily condition when it is at rest, when it is not in need of any sustenance or repairs, and when it is not threatened in any way. The areas of the brain Damasio and Carvalho identify are those whose job is to bring the organism as close as possible to homeostasis at any given time: regions that dictate the general movements of the body and the activity of its internal components. In humans, these brain regions include areas in the brainstem as well as the nucleus accumbens, ventral striatum, ventral pallidum, "and other basal ganglia and basal forebrain sectors" (Damasio & Carvalho, 2013, 145-146). (At the microscopic level, the authors have things to say about

unmyelinated axons in the interoceptive systems, which function to convey signals from humoral and visceral aspects of the body towards the brainstem (Damasio & Carvalho, 2013, 148-149). Although their description is an interesting and well-informed hypothesis, it is too far removed from our present topic to warrant thorough discussion here.)

Given that these macroscopic brain regions are phylogenetically ancient compared to neocortical regions in humans, and given that in human brains such “phylogenetically recent sectors [as those of the neocortex]...contribute to but are not essential for the emergence of feelings,” the authors speculate that feelings are not exclusive to humans or even mammals (Damasio & Carvalho, 2013, 143). Possession of those phylogenetically ancient regions, or their analogs, theoretically suffices for an animal to experience feelings invoked by interoceptive activity.

Along with requisite homeostasis-relevant neural structures, an organism capable of experiencing feelings must also possess the kind of body that can carry out what the authors call *action programmes*, which are the occasional causes of feelings. Action programmes are “biologically pre-set and largely stereotypical” physiological mechanisms (Damasio & Carvalho, 2013, 145). An example would be the automatic retraction of the hand from a hot stove; this can be initiated before the subject would feel pain in her hand. Certain kinds of action programmes constitute what Damasio (1994; cf. Damasio & Carvalho, 2013, 145) defines as emotions. For example, the action programme they identify as fear involves “a concert of responses” to a stimulus, including heart rate changes, flight-or-fight behaviors, and changes in attentional behaviors—in other words, the kinds of changes we Jamesians would describe as an *emotional reaction*.

In generating feelings correlating to emotional reactions, our bodies inform us of our deviations from homeostasis. Feelings play crucial roles in associative learning and gaining control over our behaviors. Felt experiences allow an organism to learn about the conditions of homeostatic imbalances and their corrections, and to anticipate future adverse or

favorable conditions. Thus, feelings function to “provide an additional level of regulation of behaviour” (Damasio & Carvalho, 2013, 143).

The effect of feelings in consciousness—that is, over and above the level of physiological description—is their presentation of valence: “the direction, positive or negative, and the intensity of the homeostatic deviations proxied by feelings” (Damasio & Carvalho, 2013, 150). The authors deem valence a necessary feature of body states. That is, any given body state is necessarily good or bad from the point of view of homeostasis (Damasio & Carvalho, 2013, 145). This aspect of valence is not intrinsic to exteroceptive mental states (e.g., perceptions of external world features). As mentioned above, exteroception does not involve feelings, but can trigger them. (Interoception, again, *does* involve feeling.) This position is reflected in Damasio’s (1994; 1999) general position that emotional feelings aid us in evaluating our situations.

In summary, Damasio and Carvalho take feelings to be mental correlates of bodily changes that inform the organism of a deviation from homeostasis. All feelings are valenced, which puts them in contrast with perceptions (of the extrabodily world). To reconstruct their claim in our preferred vocabulary: emotional experiences are a breed of affect directly caused by the physiological changes of emotional reactions. Exteroception can trigger emotional feelings indirectly, but only by way of interoceptive processes.

Damasio and Carvalho’s view is helpful here due to its emphasis on valence as a necessary component of feelings. As we will see later, valence plays a key role in the operations of emotion experiences. However, their view of experiential consciousness is restrictive: if only creatures with limbic systems or similar, massive neural structures are conscious—that is, have affectivity—then many, perhaps the majority, of the world’s organisms lack access to emotional experiences even if they demonstrate emotional reactions. Fruit flies and marine snails, for instance, show fear responses and can be conditioned to enact them in novel scenarios (LeDoux, 1996, 147; 2002a, 63-64). But if Damasio and Carvalho are correct, then

those beings' emotional reactions produce no affective correlates—no emotional experiences. Some theorists are less miserly with their attribution of affectivity throughout the biological taxa.

3.1.2.2 Ginsburg and Jablonka's evolutionary account

A promising theory of the evolutionary emergence of affect (as experience) has recently been proposed by Simona Ginsburg and Eva Jablonka (2007a; 2007b; 2010a). According to their theory, affective states are the product of particular kinds of activity across a highly interconnected system of neurons. If a creature's internal or external sensors are activated, then they will trigger a chain reaction of activity in downstream neurons throughout the system, allowing for rapid communication of sensory signals throughout the body of the organism (Ginsburg & Jablonka, 2007a, 220). The overall state of activity across the organism's neural system at a given time is one of "incessant and persistent neural stimulation of the animal's external and internal sensors"; the authors call this state the organism's *overall sensation* (Ginsburg & Jablonka, 2007a, 220). However, this overall sensation is not—not yet—equivalent to a phenomenological experience. Instead, it is "a weak, completely functionless, and meaningless side-effect of an interconnected sensory-motor system" (Ginsburg & Jablonka, 2007a, 220).

The overall sensation looks very similar to a preconscious state of primordial affectivity such as is seen in Colombetti's (2014) enactivist theory. However, a key difference is that Colombetti takes primordial affectivity to be intrinsically functional, given that it operates to make sense of the organism's perceivable world. Allowing for some leniency of definition, it looks like Colombetti's theory sees all such phenomena as cognitive—where *cognition* is here understood as any means of an organism's interacting with and navigating through its environment. Ginsburg and Jablonka's notion of overall sensation, on the other hand, is necessarily noncognitive since it is functionless.

How does an overall sensation become an experiential state, an affect? According to Ginsburg and Jablonka's model, experience comes into the picture when an organism has, or perhaps has the capacity for, several different kinds of overall sensations that follow bodily changes, and when these numerous overall sensations become integrated and persist (Ginsburg & Jablonka, 2007a, 221). Transitions between distinct overall sensations reflect reflexive sensory-motor behaviors: recoiling from objects that touch the organism's membrane, or releasing waste from the body, or other such behaviors might be modelled according to such transitions. But since the scope of phenomenality in such cases is going to be severely limited, the authors refer to such instances as *limited experiencing*: "Experiencing is limited both because it is based on a limited learning ability, which can only modulate preexisting reflexes, and because the inner feelings do not yet have a function, so evolution based on them is limited" (Ginsburg & Jablonka, 2007a, 221).

States of limited experiencing might qualify as affects if we take them to be raw feelings. It is, I think, fair to think of limited experiencing as the most rudimentary form of *what-it-is-likeness*, in the spirit of Thomas Nagel's famous posit (Nagel, 1974). But since it is functionless, if we call such a state affective, we cannot extend it to apply to emotional experience, which I have maintained has certain functions. So, we should look at what Ginsburg and Jablonka have to say about the evolution of *functional* affect or what they call *unlimited experiencing*:

A system can be defined as one with "*unlimited experiencing*" when a theoretically very large number of whole-organism sensory states can be generated, and when these states can be given a value and function as motivational states. An animal with such a system can be said to experience and feel: it can be said to have affective states and basic consciousness. (Ginsburg & Jablonka, 2007a, 221; original emphasis)

The difference maker that extends an organism's experiential capacity into one of unlim-

ited experiencing is in that organism's ability to form novel associative pathways between interoceptors or exteroceptors and patterns of neural activity that lead to behavioral changes. When such associations are made, the animal becomes significantly less constrained in its means of interacting with the world: it is not limited to innate, rigid reflexive reactions, but instead can roam freely through its environment by learning and producing "new adaptive behaviors, based on partial cues related to its idiosyncratic, individual, ontogenetic learning history" (Ginsburg & Jablonka, 2007a, 221). In other words, unlimited experiencing co-occurs with the capacity for learning novel behaviors and associating them with interoceptive or exteroceptive activity.

The authors offer a very general notion of what counts as learning (Ginsburg & Jablonka, 2007a, 222). First, an input (either interoceptive or exteroceptive) leads to a reaction that has some functional effect. Next, a physical trace of the input–effect relation persists across the neural system. Finally, the relation trace can be recalled—that is, retriggered—upon exposure to the same type of input; and that trace can be more readily recalled upon subsequent re-exposures. Importantly, the co-occurring sensory stimulation and memory traces both contribute to the overall sensory state of the organism (Ginsburg & Jablonka, 2010a, 112). This is the driving factor behind the qualitative character of the creature's experience: "At each moment, the sum-total of neural activity in the animal's nervous system, due to persistent stimuli and the activation of memory traces which have a minimal temporal, present-extending duration, is an overall sensory state — a dynamic state with a specific sensory flavour or signature, an experiencing" (Ginsburg & Jablonka, 2010a, 112).

To summarize Ginsburg and Jablonka's theory: any creature whose neural architecture allows for associative learning—*viz.* the pairing of novel inputs from interoceptors or exteroceptors with behavioral changes—is capable of experiencing affect.

3.1.3 Functions of affect

Although theorists diverge over the details of affect, they typically agree that it provides something beneficial to the well-being of its organisms. These functions can be clustered into three types that—coincidentally, of course—also inform the desiderata of a theory of emotion: informational content, intentionality, and motivation.

3.1.3.1 Affect's informational content

The majority position in the literature is that affective qualities represent current conditions of the body (e.g., Barrett, 2017; Damasio & Carvalho, 2013; Russell, 1980, 2003). A phrase often used to capture what a change in affect represents is *deviation from homeostasis*. Deviations from homeostasis often indicate detrimental changes to the organism's body such as tissue damage or a lack of metabolic fuel, though they can also indicate beneficial events such as the acquisition of fuel or the opportunity to copulate.

A shift from one point in the affective state space to another reflects a shift in the organism's bodily condition. The affective quality of a state of hunger or a headache might creep into consciousness very slowly, over the course of hours, while other affective states such as those accompanying or comprising sharp pains and orgasms might reflect instantaneous deviations from a prior state.

As mentioned above, affect is generated by interoception. But Barrett takes care to explicate her belief that interoception does not exist for the sake of producing affect (Barrett, 2017, 72-73). Nor does she believe that an explanation of interoception brings with it an explanation of affect:

Interoception did not evolve for you to have feelings but to regulate your body budget. It helps your brain track your temperature, how much glucose you are using, whether you have any tissue damage, whether your heart is pounding, whether your muscles are stretching, and other bodily conditions, all at the same

time. Your affective feelings of pleasure and displeasure, and calmness and agitation, are simple summaries of your budgetary state. Are you flush? Are you overdrawn? Do you need a deposit, and if so, how desperately? (Barrett, 2017, 73)

Barrett's position is echoed by many other theorists, and I think it's a fine working theory of affect's informational content. The general affective system informs the experiencing organism of how it is faring at the moment.

3.1.3.2 The intentionality of affect

By representing deviations from homeostasis, affective states intend toward the organism's body. The information they convey regarding the organism's body point towards *that* organism as an individual.

But on their own, affective states do not say anything about the world outside the organism. There's no self-other divide denoted by general affective feelings, and there's no indication of anything outside of the organism's body to which it can react or upon which it can act. A general affective state, uncoupled from sensory registration or perceptual states, does not point the organism toward any feature of its environment.

I should point out that not all theorists agree with this claim. Colombetti, for instance, takes the intentionality of affectivity to be world-directed. Affectivity, on her theory, is an organism's means of "making sense" of its world: "... all living systems are *sense-making* systems, namely..., they inhabit a world that is significant for them, a world that they themselves enact or bring forth as the correlate of their needs and concerns" (Colombetti, 2014, 2). She says that affectivity is the "bringing forth" of this world (2014, 21). But this construal requires that all organisms that are capable of affectivity are simultaneously capable of receiving sensory inputs from the world. An organism's affectivity, on Colombetti's view, logically depends on its possessing exteroceptive capacities. If Damasio and Carvalho are

correct in claiming that interoceptors evolved prior to exteroceptors (Damasio & Carvalho, 2013, 145), and if affect is a consequence of interoception, then Colombetti's requirement here is unwarranted.

However, if affect is intended toward the organism rather than its environment, a general capacity for affect could exist in a solipsistic creature. It is conceivable that a simple organism lacking exteroceptive systems would still be capable of experiencing affect if it possessed a capacity for interoception. Although its affective states would intend towards itself, this intentional quality would not be particularly helpful in many environments. Perhaps such technically senseless creatures existed prior to the Cambrian Explosion. Imagine a primitive bivalve creature that opens and closes a mouth-like structure in order to filter the water for nutrients, but that lacks the capacity to do anything else: it spends its life adhered to a single rock on the sea floor. Imagine also that this particular breed of bivalve has the capacity for associative learning. If the bivalve is able to form associations between its interoceptive states and its behaviors (opening and closing its "mouth"), then chances are it is also able to experience affective states like hunger and satiation. It could then learn to open its "mouth" when hungry. The gradual transition the bivalve experiences, specifically the graduation from experiencing hunger to experiencing satiation, constitutes its affective system's state space. But when it reaches the point of satiation, it may learn that leaving its mouth open leads to an unpleasant experience (due to having consumed more than it can comfortably digest). Consequently, it will learn to associate the experience of satiation with the behavior of closing its "mouth". Affectivity, then, helps this solipsistic bivalve make sense of its own bodily states and behavioral capabilities.

3.1.3.3 Affect motivates action

If Ginsburg and Jablonka's account of the minimal conditions for affective experience is close to the truth then many other creatures throughout the world's environments are capable of

at least some affective states—even creatures whose neural architectures are significantly less sophisticated than mammals or birds or octopuses. In that case, a general affect system is arguably the most primitive motivator to have evolved in Earth's organisms. How does activity across a network of neurons come to motivate its host organism?

The explanation Ginsburg and Jablonka offer is tied to their description of associative learning, and it centers around the notion of organismic value. They write that the sort of open-ended associative learning described above involves “attributing intrinsic, whole-organism ‘value’” to novel associations (Ginsburg & Jablonka, 2007b, 232). Value, as they use the term here, appears to track ecological concerns:

...the individual organism must have internal criteria for deciding whether a new association or a new behavior is appropriate: it must be able to evaluate, during its own lifetime, whether a response is generally beneficial or detrimental. It must have an internal, flexible, yet robust, evaluative system, which can assess *new* stimuli and responses in a highly context-dependent, ontogeny-sensitive manner. (Ginsburg & Jablonka, 2007b, 232; original emphasis)

If an organism is able to make such assessments of its novel input–response associations, the authors say, that organism has acquired a *reward system* (Ginsburg & Jablonka, 2007b, 232). Each such system has a value that is either “positive” or “negative” (these quotation marks are the authors’)—a quality of valence, or something like it at least. If an organism behaves in a way that leads to tissue damage (say, pushing against a sharp-ended object), this will alter its overall sensation and produce a negative affective state which will in turn lead to a different behavior which, hopefully for the organism, will lead to a positive affective state (Ginsburg & Jablonka, 2007b, 236-237). Thus, the overall function of these affective reward systems “is to inform the animal about its present deviation” from its ideal condition, “and guide it towards reaching it by directing adaptive behaviour that is based on past history” (Ginsburg & Jablonka, 2010a, 117). Once such feelings evolved in organisms, “they

became central organizing causes and navigators for the animals [*sic*] goal-oriented actions” (Ginsburg & Jablonka, 2010a, 117; emphasis removed).

While Ginsburg and Jablonka’s theory allows attribution of affect to creatures as simple as the solipsistic bivalve mentioned above,²⁹ most affective scientists are interested primarily in human psychology. Indeed, affect motivates differently according to organisms’ capacities to react to their affective states. Regarding human affect, Barrett holds that it functions not to motivate us to particular action, but to prompt the brain to search for explanations of what’s going on: that is, to make sense of one’s interoceptive state: “Your brain constantly uses past experience to predict which objects and events will impact your body budget, changing your affect” (Barrett, 2017, 73). On the other hand, to seek and acquire knowledge arguably is a form of behavior. While our solipsistic bivalve knows only to open its “mouth” in response to its feeling of hunger, we humans can train ourselves to recognize our hunger as indicative of a need to *acquire food*. We can then act on that need with the intention of acquiring food—not simply (like the bivalve) with the intention of diminishing our feelings of hunger.

3.1.4 Emotion is not merely affect

With these three related functions, affect is capable of helping an organism react to its bodily requirements as they are made salient. A general affect system thus plays a crucial ecological role for any creature that has one.

Ginsburg and Jablonka’s portrayal of affective states as motivational experiences gets us a good distance toward what we need in order to ground an explanation of emotional experiences. Their theory offers a nice middle ground between the theories described previously: although it agrees with Damasio and Carvalho’s claim that affective states are intrinsically

²⁹In fact, Ginsburg and Jablonka would not think it likely that our solipsistic bivalve is capable of “full experience”, since they take both interoception and exteroception as necessary contributors to the process. But their reason for this commitment is theoretically shallow: “there is no point in being aware of one’s internal states if one cannot do something about it” (Ginsburg & Jablonka, 2010a, 121). They do not offer an argument for this commitment other than deferring to other sources. Thus, their commitment is not a consequence of the theory of experiencing that they present, so we have no reason to follow them on that point.

valenced (and thus motivational), it extends drastically the range of creatures that might have access to affective states. And it achieves this while maintaining that affect is necessarily qualitative.

Going forward, I will assume that Ginsburg and Jablonka's portrayal of affect is correct, or at least as close to correct as we'll get right now. That is, I'll take an affective state to be a valenced "raw feeling" (as Barrett says) that functions to motivate its organism to behave in self-benefitting manners. Negatively valenced affects are ideally self-defeating: unpleasant feelings motivate their organisms to try to rid themselves of such feelings. And positively valenced affects are ideally self-maintaining: pleasant feelings motivate their organisms to continue behaving in such a way as to keep such feelings around as long as possible.

On their own, however, general affective states are insufficient to qualify as emotions. Those who adhere to a cursory, standard reading of James's theory of emotion will likely see my claim here as a departure from it: contrarily, they will say, James simply takes emotions to be affects and nothing more. But, as Chapter 1 demonstrated, Jamesian emotions are not the psychologically uninteresting epiphenomena many readers take them to be. Instead, Jamesian emotions meet the three desiderata demanded of a theory of emotions: that they are motivational, information-bearing, and world-directed intentional mental phenomena. We've just now seen that *all* affective states are motivational, primarily because they are intrinsically valenced. If emotional experiences are breeds (or a breed) of affective states, then they get one of the three desiderata for free. But emotions, I will argue, carry different kinds of information, and this information is intended towards both the organism *and* its external environment. While all affective states motivate their organisms to act in ways determined solely by internal factors (i.e., deviations from homeostasis), emotions are unique in that they motivate their organisms to *react* to the world as presented to them. The remainder of this chapter details these unique qualities and functions of emotional experiences.

3.2 More than affect: emotion and perception

At the end of the previous chapter, I posed a question: Can emotional reactions be triggered by nonmental stimuli? I want to argue, against the general consensus, that they can. However, I need to avoid thereby reducing emotions to mere affects. I have promised a story according to which emotional experiences are a special kind of affective state, differing from other affects according to the information they carry and the intentionality of that information.

In asking what sort of information emotional experiences bear, we simultaneously ask what that information is about. Damasio, according to the summaries in this and the previous chapter, runs with one answer: emotional feelings inform the organism about the state of its body in connection with a stimulus object. Similarly, Prinz describes the informational content of emotional feelings as the condition of the body, but also emphasizes that the feeling conveys the body's preparedness for action. On both accounts, the occasional cause of the emotion is the perceptual state that triggered the emotional reaction, while the bodily changes constitutive of the emotional reaction are the direct cause of the emotional experience. Emotional experiences provide us with information about both the external world and our own bodies.

Peter Goldie (2000; 2002; 2009) similarly states that emotional experiences ("emotional feelings") involve both kinds of intentionality. He claims that "emotional episodes" involve two distinct kinds of feelings running in tandem: bodily feelings, roughly of the Jamesian sort, along with what he calls *feelings towards*. These latter feelings are intentionally directed at the object of emotion: one can have an angry feeling toward one's rival, for example, or a lustful feeling toward a potential mate. This, he says, is a departure from the Jamesian tradition (and, in fact, from much of the cognitivist camp) which allows only for bodily feelings.

Consider from the first-person perspective James's famous example of reacting with fear

towards a bear. In this case, the bear, a real-world entity, is the stimulus that triggers your emotional response (say, gasping, freezing, and all the bodily changes required to make that happen), and you experience the subsequent affective state as a feeling of fear. Now, it is perfectly reasonable to add to your description the claim that you are afraid of the bear. Therefore, if we are going to fully explain our emotional experiences, we need to make room in our explanation for this intentional aspect. But notice that we have not established that the emotional stimulus—the real-world bear you’ve encountered—is the very same as the object towards which you feel afraid, or the *object of emotion*. The object of your emotion in this case may not in fact perfectly match the bear as it is in the real world, but rather your own interpretation of what you perceive. In some cases, we react with wild exaggeration towards perfectly harmless or ecologically neutral stimulus objects. In episodes of objectless emotions we react to... nothing—not an object present to consciousness, at least. So, if there are cases in which the occasional stimulus is not identical to the object of emotion, then we have no reason to think that the intentional quality of an emotion necessarily enters by way of perceiving the occasional stimulus.

But intentional information about the external world makes its way into our emotional experiences somehow. In fact, such intentionality, I will argue, works its way into perceptual states *by virtue of* logically prior emotional experiences. If this information does not enter through either a perceptual state or a cognitive appraisal, how does it get into emotional experiences?

The best way to answer this question, I think, is to look backwards in time, and then work backwards in process. We know that perceptual contents are intentional, information-bearing states. If emotional experiences contribute to perceptual states, an analysis of perception can help us understand how emotional experiences satisfy the three desiderata. Therefore, the next subsection (3.2.1) looks at the influences of the general affective system on the evolution of perceptual systems. The following subsection (3.2.2) discusses the the-

ory of affordances with respect to perceptual states. Following that, subsection 3.2.3 explains how emotional experiences partly constitute those perceptual affordances.

3.2.1 Affect and the evolution of perceptual systems

We can understand the interactions between emotion and perception by considering the evolution of perceptual systems. Researchers often point to the Cambrian Explosion as the era during which sensory systems became specialized resources for creatures to navigate and interact with their environments (e.g., Ginsburg & Jablonka, 2010b; Godfrey-Smith, 2016a,b; Trestman, 2013). Michael Trestman argues that creatures whose bodies were becoming increasingly complex—with articulable limbs and capacities for swift mobility—also required a “basic cognitive toolkit for embodied, object-oriented, spatial cognition”; he calls this *Basic Cognitive Embodiment* (BCE) (Trestman, 2013, 80). Different creatures, of course, have developed multiple discrete sensory systems: we humans and our common pets are familiar with our visual, auditory, tactile, olfactory, and gustatory senses; bats and porpoises employ distinct mechanisms of echolocation; sharks, rays, and electric eels pick up on electrical fields;³⁰ and so on. How these various specialized sensory systems developed is a broad and ongoing area of research that spans numerous disciplines. However, philosophers and psychologists enjoy explaining things in terms of general functions. Is there a functional account that can help us to understand how perceptual systems in general came to be?

One such theory is proposed by T. G. R. Bower (1974). According to Bower’s theory, early creatures possessed a perceptual system that picked up on information through a single modality, probably tactility. This primitive perceptual system then guided the evolution of specialized sensory systems such as eyes and ears. While “each sense organ is specialized to pick a particular band of stimulation” (Bower, 1974, 141), such as the visual and auditory

³⁰Add monotremes (echidnas and platypuses) to the list of electrolocating animals, since a duck-billed, otter-footed, beaver-tailed, poison-spurred, egg-laying, lamprey-eyed, milk-sweating mammal would be boring without that extra feature.

spectra, they all derived from and contribute to the same object-oriented perceptual system.

Crucially, then, the effective stimulus for any given sensory system need not be understood as the energy or medium involved in the physiology of sensory receptors. Take the visual system for instance. Bower writes that it “does not respond to light at all, it responds rather to edges, changes, ratios, and relations” (Bower, 1974, 142). This follows from the idea that for the visual system to guide behavior and decision-making in an organism, it requires complex inputs, “and *these inputs can be specified without reference to light*” (Bower, 1974, 143, original emphasis). So, while the eye is sensitive to changes in light, the significance of this sensitivity to the organism’s overall perceptual system is determined by the information carried by changes in the intensity and contrast of light. Of course, this information specifies the spatial configuration (and changes therein) of the environment surrounding the organism. A dark shadow in an otherwise light visual space might indicate a predator or prey or a potential mate.

Such sensory systems would work in service of BCE: they would allow an organism to navigate its world and consider what can potentially impact it from outside of its own body. Compare an organism with this capability to the aforementioned solipsistic bivalve (section 3.1.3.2). Since the bivalve lacks exteroceptors of any kind (including tactility), it lacks any means of picking up on information regarding its environment. It could not react to its surroundings, but only to its internal states (as represented by its general affect system). But once such a creature develops a primitive perceptual system, it can begin to react to the world that impacts it. With a tactility system in place, it can respond to changes in pressure on its body (perhaps another rock or larger bottom-dwelling organism came to rest on top of it). With a rudimentary visual system (such as found on some extant bivalves (see Morton, 2008)), it can respond to looming stimuli that may threaten it—even if it cannot distinguish between a predator and a passive danger such as a falling rock.

Perhaps there’s a psychological point here, too. The story of the solipsistic bivalve saw

it endowed with a general affective system as the most rudimentary form of consciousness. Regardless of its inability to perceive anything outside its body, it is capable of a limited range of experiences, namely states between hunger and satiation. If its descendants then developed the kinds of exteroceptors required for a primitive perceptual system, those exteroceptors would contribute inputs to the general affect system. In this new and improved bivalve, a general affective system would respond to the inputs of a primitive perceptual system. Feelings would arise as internal reactions to external stimuli. These feelings would direct behavioral responses which in turn would present new or altered stimuli to the organism as it behaves and moves within its environment. While the solipsistic bivalve merely acts upon the world, its perceptive descendant *reacts to it*.

The larger point here is that the general affect program guides the evolution of sensory systems by informing an organism of the state of its body with respect to external stimuli (registered by those systems) and motivating certain responsive behaviors. In ecological terms, the affect system makes salient the significance of certain stimuli. Stimuli are rendered salient because they stand in some relationship to the organism's bodily well-being. This puts pressure on the organism's descendants to maximize their sensitivity to that stimulus kind. The greater sensitivity to a stimulus kind, the greater the chance an organism can stay alive and help its lineage prosper. And so, features of the world are recognized as to-be-responded-to. As certain responsive behaviors lead to the propagation of the individual, and thereafter the species, the sensory mechanisms that trigger those behaviors are retained and refined throughout generations.

Later into this chapter I will argue that the affective qualities that accompany sensory inputs constitute emotional experiences, or emotions proper. In order to set up that argument, we need to further examine the functional nature of perception.

3.2.2 Affordances

The intentionality of emotional experiences and the kinds of information they provide ultimately can be explained in the vocabulary of James J. Gibson's ecological psychology (Gibson, 1966, 1979). Gibson is best known for his theory of perception: the contents of perceptual states are what he calls *affordances*. In the remainder of this chapter I will utilize a broadly Gibsonian framework, but I should note that I am not undertaking a project of Gibson scholarship. Whether the notion of affordances I advocate below is entirely faithful of Gibson's theory is orthogonal to whether concepts derived from his work can help us understand the proposed primitivist theory of emotion.

Here is Gibson's introductory description of the notion of affordances:

The *affordances* of the environment are what it *offers* to the animal, what it *provides* or *furnishes*, either for good or ill. The verb *to afford* is found in the dictionary, but the noun *affordance* is not. I have made it up. I mean by it something that refers to both the environment and the animal in a way that no existing term does. It implies the complementarity of the animal and the environment. (Gibson, 1979, 127; original emphasis)

Affordances are found throughout the natural and artificial world. A common case involving functional human artifacts might help us to understand them. Consider the following fact:

F₁ Ladders are for climbing.

We can restate this fact in terms of what such artifacts allow us to do with them:

F₂ A ladder affords climbing.

I expect you're still on board with this. We can restate this once again to ascribe certain properties to the ladder according to Gibson's preferred terminology of *affordances*:

A₁ A ladder *is climb-up-able*,

where *is climb-up-able* describes the ladder's affordance with respect to the behavioral capacities of an organism (see Gibson, 1979, 128). Affordances depend on both the objects and the organisms (potentially) interacting with them: ladders afford climbing to those organisms who can climb (humans, other primates, maybe cats, mice, bears...) but not those with drastically different anatomies (deer, sharks, jellyfish...). Gary Hatfield gives an analogous example: wood is nutritious to termites but not humans (Hatfield, 2009a, 292). In affordances-talk, wood affords nourishment only to those organisms (e.g., termites) whose metabolisms can process wood.

Affordances are often said to exist independently of whether or not they are in fact perceived (Scarantino, 2003, 954; see also Chemero, 2009, ch. 7). A ladder in an otherwise unoccupied shed is *climb-up-able* in the abstract; wood outside the vicinity of termites is *metabolize-able* to termites in the abstract. If so understood, then all we need to explain the ladder's *climb-up-ability* (the conditions under which that object affords an organism some interactive behavior) is a description of the ladder's structure, a description of the organism's physiology and anatomy, an observation of the organism's sensitivity to the affordance, and perhaps some spatiotemporal commonalities between ladder and organism. Affordances may then be a kind of disposition or potentiality, and organisms become aware of them when perceiving the ladder. Perceiving an affordance, on this view, means perceiving an object's "opportunities for behavior"—such as the opportunity to climb up the ladder—in an abstract sense (Chemero, 2009, 135). Human perceivers can consciously entertain affordances and choose whether or not to act upon them. Nonetheless, we can also appeal to numerous affordances that are not perceived to describe why certain states of affairs come about. A chair leg affords me *toe-stub-on-ability* when I'm walking around in the dark or with my eyes closed.

But notice something important here regarding the distinction between accidental movements and intentional behaviors: the ladder *can* be climbed only if the perceiver *per-*

ceives it as climb-up-able and is able to act according to that perception. And perceiving, as Gibson (1976) exclaims, is very much an activity, not a passive event. Conversely, the organism cannot climb the ladder if it does not actively register that the ladder affords climbing (that is, if it does not see or feel or otherwise sense the ladder).³¹ There is, here, a very important sense in which the perceived affordance itself is constituted, in part, by the very act of perception. To put it another way, the directing of the organism's sensory system towards the ladder is a constituent of the perceived affordance just as is the size, shape, and solidity of the ladder itself. The perceiver is an actor; the activity of climbing is afforded not just by the ladder, but also by the perceiver's control over its limbs and digits. Its ability to coordinate the movements of its limbs and digits is assumed, of course. If it is unable to actively raise one leg, then the other, while gripping the sides of the ladder, then that ladder does not afford *climb-up-ability* relative to its perceiver. But consider, now, that just as the coordinated movements of the limbs and digits constitute behavior, so too does directing one's sensory systems (automatically or deliberately) through one's environment. In perceiving the object that is the ladder, the perceiver itself contributes to the structuring of the relationship that constitutes the affordance of the ladder's *climb-up-ability* by that perceiver.

The qualification that affordances for intentional behavior are partly constituted by an act of sensory registration allows us to separate these affordances from the dispositional, abstract affordances described by the likes of Andrea Scarantino (2003). Let's consider the accidental or dispositional sense of affordance Φ -*affordances*. These can be given broadly physical descriptions (the ladder's structure, the organism's physiology and sensitivity, and spatiotemporal relations between the two). Those affordances perceived through acts of sensory registration we can call Ψ -*affordances*. It is these latter affordances that constitute

³¹I discount from discussion those wildly implausible "gotcha!" thought experiments concocted by certain breeds of analytic philosopher. While it may be *physically* possible that a nonperceiving organism's limbs could individually spasm in such a way that their random movements would result in that organism's climbing the ladder, such an event is *practically* just about impossible, and theoretically uninteresting. Appeal to such a counterexample would merely afford its utterer *disrespect-ability*.

the contents of perceptual states. Their designation as Ψ -affordances respects their status as members of a psychological ontology.

Ψ -affordances are not merely psychological analogues of Φ -affordances. They are a specific breed of affordance that causally depend on not only organisms' motor capacities, but their *sensory*-motor capacities. I believe this is the best way to understand Gibson's following quote, which causes much drama throughout the literature on affordances:

An important fact about the affordances of the environment is that they are in a sense objective, real, and physical, unlike values and meanings, which are often supposed to be subjective, phenomenal, and mental. But, actually, an affordance is neither an objective property nor a subjective property; or it is both, if you like. An affordance cuts across the dichotomy of subjective–objective and helps us to understand its inadequacy. It is equally a fact of the environment and a fact of behavior. It is both physical and psychical, yet neither. An affordance points both ways, to the environment and to the observer. (Gibson, 1979, 129)

Relative to Ψ -affordances, it is much more difficult (for me, at least) to understand how a perceiver-independent Φ -affordance cuts across the subjective-objective dichotomy. This reciprocity between perceiver and object in *determining* perceptual content (Ψ -affordances) is crucial for understanding the sense in which perception of these affordances is *direct* rather than representational: the perceptual content *is* the set of events that can occur when the perceiver behaviorally interacts with the object. It makes little sense, then, to say in isolation of a perceptual state that *the ladder affords climb-up-ability*. This Gibsonian phrase should be taken as short-hand for the longer description of the Ψ -affordance:

A₂ The ladder affords the perceiver the ability to climb if and only if the perceiver succeeds in picking up (*viz.* sensing) certain indicatory information emitted by the ladder.

We can identify, then, three different sets of variables that determine an object's Ψ -affordance

(e.g., a ladder's *climb-up-ability*) *qua* perceptual content: the physical structure of the object (measured in units relative to the organism's body size: see Gibson, 1979, 127-128); the sets of patterned movements and physiological changes (behaviors) the organism is capable of undergoing; and the act of sensory accommodation and registration (e.g., gazing at, sniffing at, touching the object) the organism does, in fact, undergo. And we can express a Ψ -affordance using the following formal notation:

$$P \vec{a} B_O$$

Here, P marks out the perceiving organism (with its particular physiological features and capacities), \vec{a} is an affordance to act out a behavior, B, with respect to some object O. The arrow above the *a* designates that the affordance is one of the perceiver P acting towards the environmental object O. Thus, A_2 becomes:

$$A_3 \quad P \vec{a} \textit{climb-up-ability}_{\textit{ladder}}$$

Note that the P in a Ψ -affordance will be experienced from the first-person perspective. The above notation does not describe a propositional statement with a necessarily explicable concept of self; it is intended as an approximate, symbolic depiction of an experiential state. Indeed, it is intended to capture a kind of mental state available to any creature with the kind of neural architecture described by Ginsburg and Jablonka's theory (see section 3.1.2.2 above). Such mental states are unlikely to qualify as *cognitive*, at least according to advocates of the cognitivist approach described in the introduction to the dissertation and in section 2.1. This, I believe, is a strength of the proposed noncognitivist theory: one's *self* can occupy a specific perspective and take part in a relational perceptual state without depending on access to a concept of self.

With this theory of perceptual content on the table, let's now shift discussion back to emotion.

3.2.3 Emotions and Ψ -affordances

I have stated at numerous points in this chapter that emotions partly constitute perceptual states. Moreover, I have just argued that the contents of perceptual states are a special kind of *affordance*, namely Ψ -*affordances*. I now need to show how emotions partly constitute Ψ -affordances.

James Gibson does not much mention the relationship between perception and emotion in his works. However, he does mention something telling. To test hypotheses derived from his theory of ecological psychology, Eleanor Gibson and Richard Walk (1960) develop the “visual cliff” experiment. This involves placing an infant, a kitten, or some other animal at the edge of a glass floor, through which can be seen a patterned surface. Crucially, the underlying patterned surface “drops away” from the glass at a certain point. James Gibson describes the results as follows:

The animals or babies tested in this experiment would walk or crawl normally when they could both see and feel the surface but would not do so when they could only feel the surface; in the latter case, they froze, crouched, and showed signs of discomfort. ... For my part, I should feel very uncomfortable if I had to stand on a large observation platform with a transparent floor through which the ground was seen far below. (Gibson, 1979, 157)

Is the lack of comfort due to a prior perception of a negative Ψ -affordance, namely that the sudden drop of surface affords the perceiver *fall-off-ability*? Or is it instead a continuation of the affective state that helped to constitute that Ψ -affordance in the first place?

Recall the distinction between interoceptive and exteroceptive nervous activity. This was discussed above in subsection 3.1.1. Recall also the Jamesian line that bodily changes are felt as they occur, discussed at length in Chapter 1. Such changes include not just movements of the limbs and activity of the viscera, but also excitation of sensory channels during sen-

sory registration. Although exteroception itself is distinguishable from interoception (*per* Damasio and Carvalho's description: see subsection 3.1.2.1), the neural processes underlying exteroception may be represented as subject to interoception in nature, since they involve physiological changes. Or, as I put it in the previous chapter, excitation of sensory systems is just as legitimate a bodily change as is, say, an emotional reaction such as crying or laughing. The difference between these cases is not of kind, but of scale of activity. Sensory system changes are microscopic, while crying and laughing—identifiable emotional reactions—are macroscopic changes.

I propose that the affective states resulting from excitation of sensory systems during stimulus registration constitute emotional experiences, or emotions proper. As the sensory systems resonate with or accommodate to their stimuli (e.g., light patterns falling on the retinae), the neural pathways running throughout the brain (and body) trigger affective states that contribute predictive values based on associative pathways built from prior encounters with such stimuli (cf. Barrett & Bar, 2009; Ginsburg & Jablonka, 2007a,b, 2010a).

Lisa Feldman Barrett and Moshe Bar (2009) provide neurological evidence in favor of a similar hypothesis. Specifically, they argue that the orbitofrontal cortex (OFC) plays a crucial role in combining information from exteroceptive sensory systems as well as interoceptive representations to determine the recognitional content of a perceptual state:

The centrepiece of this circuitry is the orbitofrontal cortex (OFC). . . . The OFC is a heteromodal association area that integrates sensory input from the world and the body (i.e. from extra- and intrapersonal space) to create a contextually sensitive, multimodal representation of the world and its value to the person at a particular moment in time. . . . The OFC's ongoing integration of sensory information from the external world with that from the body indicates that conscious percepts are indeed intrinsically infused with affective value, so that the affective salience or significance of an object is not computed after the fact. (Barrett

& Bar, 2009, 1329)

As part of this heteromodal integration of inputs, the state of the body—represented as an affective state—contributes to the brain’s recognitional capacities. Combining this description with the Jamesian claim about felt bodily change gives us an account of specific affective states that partly constitute perceptual content. In other words, the *interoceptive* representation of excitation in *exteroceptive* sensory channels provides the *affective* component of a perceptual state. What’s needed for this to occur is an associative mechanism that can represent an association between some pattern of exteroceptive activity and a temporally nearby interoceptive state. The interoceptive state grounds an affective state that is retriggered by subsequent patterns of exteroceptive activity. And it is precisely these exteroception-triggered affective states that constitute emotions proper.

However, Barrett and Bar’s theory does not straightforwardly fit with an account of perceptual affordances. Instead, they take behavior to be mediated by object recognition. The role of affect in perception, as they see it, is to help fill in the details following a rudimentary interpretation of what is being registered by the senses:

There is accumulating evidence that during object perception, the brain quickly makes an initial prediction about the ‘gist’ of the scene or object to which visual sensations refer... [T]he brain uses low spatial frequency visual information available from the object in context to produce a rough sketch, and then begins to fill in the details using information from memory... (Barrett & Bar, 2009, 1328)

They then elaborate on the role of affect in this process of object recognition:

With gist-level visual information about the object, the medial OFC initiates the internal bodily changes that are needed to guide subsequent actions on that object in context. The ability to reach for a round object and pick it up for a bite

depends on the prediction that it is an apple and that it will enhance one's well-being in the immediate future because it has been done so in the past. (Barrett & Bar, 2009, 1329)

In other words, the affordances of the apple—its *reach-to-ability* and its *eat-ability*—depend on its first being perceived *as* an apple. The role of affect in this sequence is to provide the experience-based information (e.g., that apples are delicious) that turns the perceptual content from a 'gist' to a categorical representation.

Conversely, the Gibsonian view I described in the previous subsection takes the affordances themselves to constitute contents of perceptual states. Instead of acknowledging the apple's *eat-ability* by first recognizing it as an apple, the Gibsonian view states the reverse: that we recognize the object as an apple by first acknowledging its *eat-ability*. And I believe that the account of affective, associative learning via unlimited experiencing, provided by Ginsburg and Jablonka (see section 3.1.2.2), can help us make sense of why this would be.

Recall that, according to Ginsburg and Jablonka's theory, different affective *overall sensations* reflect different patterns of neural activity in response to sensory stimulation. Associative pathways are formed between, on the one hand, both interoceptive and exteroceptive sensor activation, and, on the other, behavioral responses tuned to either the positive or negative aspects (i.e., the valence) of that overall neural state. Different overall sensations reflect different behavioral capacities and motivate those behaviors by functioning as affective *reward systems* (Ginsburg & Jablonka, 2007b, 232).

Crucially, if this all occurs during sensory registration in creatures such as us humans, then our affective reward systems must be at work during the process of perceiving extrapersonal objects (such as apples, or ladders, or James's bear). And if our affective overall sensations experienced during acts of perceiving reflect our body's behavioral capacities, then *those affective states partly constitute the Ψ -affordances that are our perceptual contents.*

Let's look at another example of a Ψ -affordance:

A₄ P \vec{a} *eat-ability*_{apple}

Recall that P is the perspective of the observer, the embodied feeling-self (see Chapter 2, subsection 2.2.4). The stimulus object, the apple, is here perceived under an aspect: ultimately it is seen as a thing-to-be-eaten. How does this perceptual state eventuate? An emotional experience enters here as the general affective system's response to the patterned sequence of neural activity originating at the organism's retinae. That pattern is associated, through prior experience or maybe through adaptation, with representations of digestive satiation and nourishment. (In perceptually complex creatures more specific memories of pleasant taste and texture profiles undoubtedly also play a role in this operation, but such sophisticated capacities are probably unnecessary for rudimentary perception of Ψ -affordances.) Since these states are direct causes of pleasant feelings, a faint echo of a pleasant feeling constitutes a microscopic emotional experience occurring simultaneously with sustained sensory registration of the stimulus (the apple). The overall perceptual state that finds its way to the perceiver's consciousness is a conglomerate of the affective state, triggered by interoceptive monitoring of exteroceptive activity, and the particular patterning of that exteroceptive activity. The emotional experience's intentional character informs the organism of the state of its body and its most prominent or pragmatic suite of behaviors available as reactions to the stimulus. In plainer language: the emotion tells the organism that what is being sensed is something that will benefit it if it eats it. The stimulus object affords the organism *eat-ability*. And that perceived *eat-ability* is what allows perceptually sophisticated creatures such as humans to see the *eat-able* stimulus *as* an apple.

This example demonstrates that by synthesizing Ginsburg and Jablonka's theory of affect with Barrett and Bar's hypothesis regarding affect's contribution to perceptual processing, we arrive at a theory that sees a particular breed of affective state at work in determining the contents of perceptual states based on the perceiver's capacities for behavior towards the perceptual stimulus.

On what grounds, now, can we justify calling these exteroception-triggered affective states *emotions*? This is where the neo-Jamesian side of the theory is made explicit. The most important claim of James's theory of emotion is that an emotional experience is a feeling of a bodily change as it occurs. As I argued in the previous chapter, and hinted at above on page 126, it does not matter whether the bodily change is microscopic or macroscopic in nature. All that matters is that the reaction leads to a feeling of some sort. When exteroceptive neural channels are excited, they produce microscopic affective states. Thus, microscopic bodily changes result in microscopic emotional experiences.

However, not every feeling of a bodily change constitutes an emotional experience: many of our conscious experiences are merely affective states rather than emotions proper. For a feeling to qualify as an emotion (or an *emotional experience*), it must satisfy the three desiderata of a theory of emotions: it must be motivating, information-bearing, and intentional. While all affects are intrinsically motivating (see section 3.1.2.2), only those affective states involving exteroceptive neural activity carry information about the world outside of the body. Therefore, it is only this specific breed of affective state that can satisfy the two other desiderata. The affective states at work during acts of perception qualify as emotional experiences.

So much for microscopic emotions. What about those macroscopic phenomena we pre-theoretically identify as emotions (even if not as specific emotions)? These feelings often seem to be downstream consequences of our perceiving something, not (just) constituents of those perceptual states. Indeed, pretty much every other theory of emotion examined throughout this dissertation characterizes *emotion* in this macroscopic fashion. The primitivist theory has to account for them if it is to succeed as a theory of emotion.

One more example of a Ψ -affordance can help. Recall James's bear:

A_5 $P \vec{a}$ *be-hurt-by-ability*_{bear}

Just as the profile of sensory system activity from the apple is associated with feelings of

digestive satiation and nourishment, so too is the bear's profile—specifically details such as size and distance relative to P, and other cues such as a deep, loud roar—associated with feelings of physical damage. And, of course, physical damage is a direct cause of unpleasant feelings (the greater the damage, the more negatively valenced and highly aroused the affective state).³² Again, a faint echo of the unpleasant feeling of pain constitutes a microscopic emotional experience occurring simultaneously with sustained sensory registration of the stimulus (the bear). This emotional experience relays to the organism that its body is in a state of being harmed and represents the stimulus as the source of that harm. The resulting perceptual state contains *be-hurt-by-ability* with respect to the present stimulus–organism relationship.

Next, that perceptual state contributes to the generation of a macroscopic emotional experience. Subsection 3.1.3.3 argued that all affective states motivate action. Negative affective states motivate the organism to eliminate them. And so, the microscopic emotional reaction that partly constitutes this perceptual state motivates the organism to diminish its presence. It does so by participating in the activation of a sensory-motor association whose suite of potential bodily responses includes, say, quickening the heart beat, releasing adrenaline (epinephrine) into the bloodstream, maximizing the sensitivity of exteroceptors (widening the eyes and reorienting the head toward the stimulus), broadening the respiratory channels, diverting oxygenated blood to the skeletal muscles, and tensing the skeletal muscles in preparation for a swift exit from the scene. A manifestation of these bodily responses constitutes the macroscopic *emotional reaction*. This consequently is felt as a macroscopic emotional experience that may be familiar to us as *fear*. (The next chapter will have much more to say about how we recognize and interpret our emotional experiences.)

³²The affective character of these feelings might resemble some understanding of *fear*, but let's not commit to that just yet.

Conclusion

This chapter has attempted to develop a novel theory of emotion based on the major tenets of William James's famous (or infamous) theory as described in Chapter 1. Several descriptors attach themselves to the proposed theory.

Firstly, it is *noncognitivist*: it rejects the claim that all emotions require sophisticated internal mental processes (such as evaluative appraisals) as causes or constituents. Instead, emotions can be triggered through excitation of neural pathways during exteroceptive sensory registration.

Secondly, it is legitimately *neo-Jamesian*: it defines emotions as feelings (*emotional experiences*) due to bodily changes (*emotional reactions*). Thus, it is not the case that “we cry, strike, or tremble, because we are sorry, angry, or fearful”; but instead “we feel sorry because we cry, angry because we strike, [or] afraid because we tremble” (James, 1884, 190; 1890, II: 450; 1992, 352).

Thirdly, it is *ecological*: emotions inform us of our behavioral capacities insofar as they help us to navigate the world that impacts us. Emotional experiences partly constitute the psycho-physical affordances of our perceptual states.

Fourthly, it considers emotions to be *primitive psychological phenomena*. Emotions do not logically or causally depend upon other, more fundamental psychological kinds such as cognition or perception. Instead, they rely on interoceptive monitoring.

Taking these labels together, we can think of the proposed theory as a *primitivist theory of emotion*. This theory is offered as a plausible alternative to cognitivist theories. However, there is yet much to be explained in order to convince an audience leaning towards cognitivism that the primitivist theory can pull its weight. In the next chapter, we will investigate how we identify our emotional experiences through introspective reporting. We will also see whether, and to what extent, the primitivist theory can explain our more sophisticated, culturally informed emotional experiences such as guilt and shame.

Chapter 4

How to feel bad about yourself

Introduction

We humans are very expressive, social creatures. We thrive on letting others in our communities know how we are doing at the moment. The hungry, altricial infant wails pathetically until something activates her suckling reflex; she wails again when she soils herself. This general theme endures for a while, until eventually wails are phased out in favor of more coherent utterances such as “I’m hungry!” and “Oops...”. (Actually, these utterances often are accompanied by more wailing.) With luck and time, the developing individual will ultimately turn to making conversation, music (wailing sometimes counts), art, and literature as outlets for expressing her feelings.

Working in tandem with all these forms of expressiveness are our capacities for understanding others’ expressions of feelings. We pride ourselves on our abilities to interpret and sympathize with others’ mental states. Poker players and psychotherapists build their careers on these skills. In fact, working through the clues others give us is a rewarding passtime, often preferred to more obvious and direct methods. To grasp what I mean, see what you make of this verse by musical artist Rich Terfry (better known as Buck 65):

When I cheated on Sarah
There wasn't a star in the sky
It was covered in clouds
And I started to cry
I was cold and I deserved to be
Hardened and worthless
Windows painted shut
Rebel without a purpose
Clown smoking cigarettes
I think I thought I heard her name
Killing time with my bare hands
Waiting for the hurricane

These words don't directly convey Terfry's state of mind; they conjure images that his listeners, he assumes, would typically associate with feelings of sadness, regret, and guilt.³³ This is a classic function of verse and metaphor. Terfry could have said something much more obvious and direct, along the lines of "I felt guilty" or "I felt bad about myself". Granted, such concise statements wouldn't work so well alongside the somber slide guitar, stripped-down drums, and moaning strings of the backing track. Putting aesthetics aside, we linguistically capable adults consider generating a concise and explicit report of one's emotion a simple task (at least typically, though exceptions abound). In this case, Terfry has made things complicated for us by expressing his feelings indirectly. From the listener's perspective, it'd seem

³³I should note that, in fairness, the story Terfry tells in this song may well be fictitious or fictionalized. We might prefer to consider Terfry, the artist, as distinct from the character Buck 65, his musical moniker. Indeed, Terfry makes no mention of Sarah in his recent biography (Terfry, 2015). This is telling, as he dwells on numerous other relationships throughout his life. On the other hand, Terfry also warns his readers in that very biography that he's an unreliable narrator. Here's a link to the song in case you'd like to hear it: <https://youtu.be/FclZa3rdsCA>

unnecessary for us to work through clues to understand how Terfry feels in this episode had he opted for one of those concise statements. Straightforwardly, he'd feel guilty, or he'd feel bad about himself.

In fact, when exploring and explaining from a psychological standpoint the nature of emotion reports—even in their most concise forms—the picture is not at all simple or transparent. There is a wealth of literature across different disciplines strongly suggesting that introspection doesn't yield neat, direct information about our psychological states. Instead, our verbal reports of our emotions are at best approximations, at worst full-blown confabulations. The apparently simple task of saying "I feel bad about myself" requires significant metacognitive processing, conceptualization, and contextualizing. More concisely: we work through the clues we observe in ourselves to produce reports of our emotions.

This chapter explores how reporting one's emotional state likely occurs according to the primitivist theory of emotion. The primitivist theory, developed in the previous chapter, construes emotion as a functionally specific breed of affect. An emotion is not reducible to a perceptual or cognitive state. Yet our capacities for reporting on our emotional experiences undoubtedly require access to, and interpretation of, our emotion systems by our cognitive systems. The present objective is to develop a plausible account of such a process.

The previous chapters have prioritized explanations of emotional experiences. Emotional experiences are emotions proper, at least according to primitivism. But the realm of emotionality is not exhausted by the realm of emotional experiences. As Terfry's lyrics illustrate, we humans find ourselves in emotional situations involving numerous interwoven features: references to emotionally relevant entities outside of our own minds; names for our experiences; imagery that functions to convey or redescribe our psychological states to others. These temporally extended, multifaceted *emotion episodes* are the stuff of primary interest to many philosophers of emotion. In Chapter 2 I quoted Robert Solomon as a representative of this position: he emphasizes that "long-term narratives", not "short-

term neurological arousal”, comprise the subject of his inquiry (Solomon, 2003, 2). Thinkers such as Solomon needn’t concern themselves with the emotions of beasts and babies, since by definition they’ve identified uniquely human phenomena from the get-go. But there are those of us who prefer to explicitly acknowledge the extent to which we adult humans share certain psychological traits with other creatures *before* devoting ourselves to understanding our unique traits.

The chapter opens with a brief recap of the primitivist theory of emotion (section 4.1). This first section also presents some of the challenges the theory faces, namely how to account for the features articulated in our everyday emotion self-reports. The remainder of the chapter rehabilitates a restricted cognition-arousal theory: although it fails as an account of *emotion*, it succeeds in explaining *emotion self-reports*. Section 4.2 explores what kinds of affective evidence the body provides to the metacognitive system. Section 4.3 then elucidates the contributions to emotion self-reports of cognitive and perceptual factors such as emotion concepts and objects of emotion. Finally, section 4.4 offers a rough sketch of self-reflective emotion episodes such as that portrayed in Terfry’s lyrics.

4.1 The primitivist theory and its challenges

The primitivist theory construes emotion as a subspecies of affective state. Affective states are the felt products of *interoception*, the body’s internal monitoring processes. All affective states motivate organisms to act: positively valenced affects (i.e., pleasant feelings) are self-maintaining, prompting us to keep doing what we’re doing; negatively valenced affects (i.e., unpleasant feelings) are self-defeating, prompting us to act in ways that mitigate such feelings. Affective states graduate into one another, moving along the pleasant–unpleasant axis. A simple, unassociated feeling of hunger (via tightness of the stomach) is mildly unpleasant; consumption of food leads to a mildly pleasant feeling of satiation (via fullness of

the stomach); excessive consumption of food leads to a mildly unpleasant feeling of over-satiation. (The present author recalls regretting having finished every dish offered during a recent degustation outing.)

What demarcates emotion from other forms of affect is its intentional, information-bearing quality: emotion states carry information about how we can act as agents in our environments. As a subspecies of affective state, emotion states, or *emotional experiences*, also result from interoception. However, emotional experiences result from interoceptive monitoring of *exteroceptors*, namely those of the body's sensory systems that are attuned to the extraorganismic environment. When the onset of some stimulus (say, a looming shadow crossing the organism's retinae) activates associative pathways between sensory registration systems and motor response systems, the resulting affective state is not merely a reflection of the state of those associative pathways. It is about the relationship between the state of the body and the organism's environment. Though such a state does not represent to the organism what the stimulus is (this would require perceptual or perhaps conceptual representations) it does provide indications of the range of bodily responses to the stimulus the organism can undergo. It also motivates the organism towards one of those responses.

As the sensory systems resonate with or accommodate to their stimuli (e.g., as light patterns falling on the retinae cause the eye's components to alter gaze direction and focal depth), the neural pathways running throughout the brain and body trigger affective states that contribute predictive values based on associative pathways built from prior encounters with such stimuli. These externally-triggered emotional experiences contribute a first-person, affective quality to resulting perceptual states. A perceptual state thus involves in its phenomenal character intrinsic, affectively realized significance to the organism.

In the previous chapter I argued that typical states of object perception have *affordances* as their contents. Affordances, recall, are what the environment affords an organism, "what it *provides* or *furnishes*, either for good or ill" (Gibson, 1979, 127, original emphasis). We

humans see apples as *eat-able*, ladders as *climb-up-able*, floors as *walk-on-able*. But, I argued, we don't perceive these affordances unless we train our sensory systems to attend to the affordance-granting object occurrently. So, as I put it previously, a ladder affords the perceiver the ability to climb if and only if the perceiver succeeds in picking up (*viz.* sensing) certain indicatory information emitted by the ladder. The function of an instance of emotion, on this theory, is to respond to such information pick-up during sensory registration and thereby partly constitute a perceptual state whose content is a perceiver-dependent affordance.

Emotion, therefore, is a psychological primitive: it does not reduce to other mental phenomena such as perception or cognition. It is the direct result of neuronal activity, but it is not merely about the state of the organism's body. By drawing our attention to the world outside the body and motivating us to act, emotional experiences function in ways beyond the role of simpler affective states.

The central claim of this chapter is that emotional experiences constitute key pieces of evidence from which we generate self-reports of our own emotional states. As argued in earlier chapters, when searching for a definition of the term *emotion*, the emotional experience best fits the bill. The challenge this chapter must overcome is the apparent disparity between how the primitivist theory describes our emotional experiences (using formalized, theoretical language) and how we describe our emotional episodes using everyday folk language. Is it possible to respect the status of lay self-reports as genuine and informative expressions while maintaining a description of their underlying causes via the primitivist theory? That is, can we accept the primitivist theory of emotion without eliminating folk conceptions of emotions?

I think we can. Here's a synopsis of the story this chapter will tell. Following the onset of an emotional experience, our focus is on some extrapersonal object or event: the bear charging towards us, the person insulting us, and so on. At some point after the fact, we

might reflect on how we felt (and likely still feel) during the episode. To form a report of our emotion states, we check the characteristics of our emotional experiences against our stock of emotion concepts (at least, those we can recall in the moment). Simultaneously, we gather what information we can regarding our situation. Having found the best-fitting label for a situation, we report that we are experiencing, or have experienced, that specific emotion.

The first and most obvious challenge for the primitivist theory is accounting for the apparently irruptive and spontaneous nature of our emotional episodes as we identify them everyday. Although according to theory we are always in some sense emotional (except when we're not perceiving anything), everyday experience sees us enter periods of anger, joy, guilt, sadness, and the like—times when we notice that we are “being emotional”. And these episodes appear to us to be deviations from a more common, relatively “unemotional” baseline state of being. If we're always emotional, why do we deem ourselves so only sometimes?

The answer comes down to the extent to which an emotional experience enters into the foreground of consciousness. As I'll explain below, a drastic change in emotional experience, especially an increase in arousal, will flood consciousness such that we will attend to and acknowledge it: the emotional experience becomes the object of attention. And such a drastic change can be triggered by a large-scale physiological change, measurable across the body rather than merely in microscopic extrasensory pathways. Consider the jolt that runs through your skeletal muscles when you're surprised by the sudden presence of some stimulus: a peal of thunder, perhaps, or an object falling toward you. In this case, an associative pathway between your sensory system and your motor system has triggered a tensing of your skeletal muscles, comprising an *emotional reaction*. Such a drastic deviation from your immediately prior calm, unsurprised affective state is very likely to draw your attention to your situation and your new emotional experience. This marks the onset of a reportable emotion episode. By analogy, consider the corner of whatever room you're now in. Although there's

no physical break or gap between one piece of wall and another (barring shoddy construction), the sharp line that marks a 90° angled change of direction makes it easy to identify two separate walls. Think of the more drastic emotional reactions as the affective equivalent of that 90° change.

Yet other, more psychologically complicated changes may also mark the onset of an emotion episode—at least inasmuch as we identify an onset when we generate emotion self-reports. All sorts of factors can prompt us to assess our current or recent emotion states. Thus, we cannot rely only on drastic changes in our affective states to explain emotion self-reports. This fact is not lost to the Jamesian: William James's famous theory of emotions inspired a great deal of inquiry into the question of what we attend to when labelling our emotions (and those of other persons). In many cases, it seems, the emotional experience borne of physiological changes seems insufficient to inform a self-report. So, what other kinds of clues are helpful?

Chapter 2, section 2.2 discussed Schachter and Singer's attempt to update James's theory of emotion by bringing in an element of cognitive interpretation. The theory, if you recall, says that an emotion is a function of perceived bodily feelings interpreted according to an available stock of emotion concepts. The assessment of their theory was negative: Schachter and Singer hamstring themselves by proposing the cognition-arousal theory as an explanation of *emotion*. If instead we restrict its role to explanations of how we report on our emotions, far fewer issues will arise. The remainder of this chapter comprises a positive account the various phenomena, internal and external, to which we attend in generating reports of our emotion states. It thereby vindicates and updates the cognition-arousal theory as a resource for explaining those cognitively sophisticated emotion episodes that involve explicit references to a conceptual self.

4.2 Accounting for diversity of phenomenality

This section explores the diversity of feelings across different emotional episodes and how they factor into metacognitive generation of emotion self-reports. Recall the *specificity* claim from the previous section: activity in different bodily regions is associated with qualitative differences between emotional experiences. Certainly, when we reflect upon our emotional episodes we typically feel a certain way, and this particular feeling helps us to conceptualize our current state. It's in large part because I *feel* happy that I name my state *joy*. How do these feelings operate as cues for metacognition? An answer is offered in the next subsection (4.2.1). Following that, subsection 4.2.2 argues that different feelings are traceable to different bodily regions. The body is not only a producer of affect; it's also a topographical guide to differentiating between emotional experiences.

4.2.1 Russell's three primitives

Four decades after Schachter and Singer's study was published, psychologist James Russell produced a novel constructionist theory of emotion. "My hope," he writes, "is to achieve what Schachter and Singer (1962) set out to accomplish: a synthesis of (a) James's (1884) insight that emotion involves a self-perception of automatic processes with (b) modern evidence on the process involved" (Russell, 2003, 146). Such a synthesis would account for what he calls *prototypical emotional episodes*, namely those instances in which an individual acknowledges that she is in an emotional state (Russell, 2003, 146).

In setting up the conceptual framework for his theory, Russell identifies three psychologically primitive processes that, "alone or combined with information processing and behavioral planning, then account for all the myriad manifestations and influences called *emotional*" (Russell, 2003, 148). These are *core affect*, *affective quality*, and *attributed affect*.

Core affect, per Russell's theory, is raw feeling. I discussed notions of affect at length in

the previous chapter, including the two-dimensional model Russell proposes in an earlier publication (Russell, 1980). Here Russell elaborates on his model. Core affect is experienced as an irreducible mental state. It can be tracked across two dimensions: pleasure–displeasure and arousal (or activation). At any given moment, a conscious being’s affective state can be described as a point in a quadrant defined by these two dimensions. For example, “elation” might pick out a point high on both the activation and pleasure axes, while “sadness” might pick out a point closer to the opposite pole (i.e., displeasure and deactivation). Thus, for any organism, if there is something it is like to be that organism, that quality of being can be at least partly described in terms of core affect: a set of co-ordinates, for example.³⁴ Certainly as far as human minds are concerned, affect is ubiquitous: an affective state may be neutral, moderate, or extreme, and it will occupy our attention to an extent relative to its severity. Changes in core affect are felt in proportion to their distance from a prior state and the rapidity with which they occur. A change in core affect “evokes a search for its cause and therefore facilitates attention to and accessibility of like-valenced material” (Russell, 2003, 149). Finally, core affect functions to assess continuously one’s internal state.

Affective quality is a perceivable feature of all perceived objects and events (Russell, 2003, 149). It is flagged by the organism’s perceptual system during an interpretive act of perception: “Objects, events, and places (real, imagined, remembered, or anticipated) enter consciousness affectively interpreted. The perception of the affective qualities of all the stimuli typically impinging at any one time (how pleasant, unpleasant, exciting, boring, upsetting, or soothing each is) then influences subsequent reactions to those stimuli” (Russell, 2003, 149). To perceive affective qualities in an object is to notice its capacity to change core affect.

The previous chapter put forward a similar notion when discussing the emotion–perception relationship. Shifting back to the primitivist approach, affective quality becomes the feature of an instance of sensory registration that triggers an emotional reaction.

³⁴Though arguably for a creature to properly qualify as conscious under primitivism, it must undergo affective changes: see the previous chapter, section 3.1.1.

That is, it's not necessarily the bear *qua* bear that triggers a fearful emotional reaction when it enters my visual field. Rather, it's some combination of shape, size, proximity, or other ecological relationships between it and me that do the work. I need not see *that there is* some object X that possesses affective quality A in order to have an emotional reaction. I need only register the presence of affective quality A. Moreover, affective quality pick-up is *pre-perceptual*; indeed it contributes to the downstream processes that yield perception of objects. This reconceptualization of Russell's affective quality notion is inevitable for proponents of the primitivist theory.

Back to Russell's terminology. An instance of *attributed affect* sees the organism identify as the cause of its change in core affect some object or event (Russell, 2003, 149). When I perceive some object (a large, hairy spider) near my person, and I experience a change in core affect (from near-neutral calm to high arousal and displeasure), I associate my change in core affect with that object. This is roughly equivalent to attributing the folk term *frightening* to an object that frightens me.³⁵ Though I may feel that my attribution is accurate, misattribution can occur. And the function of attributed affect is to guide attention away from one's own affective state and toward its cause. Attributed affect can thereby train the perceptual system, via association, to perceive an affective quality in an object or object kind (Russell, 2003, 149).

We should not conflate affective quality and attributed affect—not on the primitivist view, at least. Again, an affective quality is picked up during sensory registration and thus triggers emotional reactions. Attributed affect, on the other hand, results from a *perceptual* (or even *post-perceptual*) process: it is the attaching of a qualitative label to a perceived object or event. But this must take place after other emotional processes have occurred: attributed affect depends upon a logically prior emotional experience with respect to some stimulus. In short, while core affect and affective quality together contribute to perceptual processes, attributed affect is an outcome of downstream perception. (Were we to carry this distinction

³⁵Where “roughly equivalent” neglects the potentially tricky differences between any conceptual resources employed either case.

into the vocabulary of philosophy of perception, we might here distinguish between an *act of perceiving* and *perceiving as*.)

Russell claims that these three primitives provide the foundations for metacognizing toward self-reports of emotion. I agree. But it's clear that they contribute in very different ways to emotion self-reports. Indeed, empirical evidence suggests that some people are more inclined to rely on internal (psychological and physiological) cues in generating emotion reports, while others turn more frequently to external (situational) cues (Laird & Bresler, 1992; Strout et al., 2004). It is plausible then that Russell's three primitives are weighted differently between these populations.

As I see things, the kinds of emotion self-reports most in need of explanation are those that derive most heavily from introspectively acknowledged changes in core affect that are triggered by the pick-up of affective qualities. Changes in core affect on their own do not suffice, as these could be brought about through purely interoceptive processes such as those that produce feelings of hunger or an urge to release bodily waste. These feelings don't qualify as emotions proper, since they don't inform the organism of potential behaviors as responses to its environment. However, the causal pairing of core affect and affective quality pick-up relates the feeling-self to the external environment in precisely the way described in Chapter 3. To restate the major claim of primitivism: emotional experiences are affective states triggered by interoceptive monitoring of exteroceptive systems.

Unfortunately, this means that reports based on attributed affect are left somewhat in the dark, at least for now. There's an important discussion to be had regarding whether, and if so how, emotion reports grounded primarily in analyses of attributed affect represent genuine emotional experiences. Such reports are more like confabulations based on situational evidence than approximations based on feelings. But since that issue is further removed from the tenets of primitivism than the kind of self-report mentioned in the previous paragraph, we must put it aside for another day.

So then: how do bodily changes and changes in core affect contribute to emotion self-reports?

4.2.2 Bodily maps of emotional feelings

According to the *specificity* claim (discussed in Chapter 2, section 2.2.1), differences in bodily change account for qualitative differences between affective states. But this does not imply that we are able to recognize introspectively which parts of our bodies are producing the emotional experiences we wish to report upon. An emotional organism might not necessarily be able to attribute its “fearful” feelings to its viscera, or certain skeletal muscles, or its respiratory system, even if activity in these regions causally accounts for said feelings. Certainly in humans the face plays a prominent role during emotion episodes: studies utilizing volumetric (Susskind et al., 2008) and thermographic (Salazar-Lopez et al., 2015) imaging methods reliably distinguish the facial contortions of different emotional reactions. So, even if we accept the *specificity* claim as fact, whether we can map our emotional experiences across the body for the sake of emotion self-reporting is yet to be determined.

Recent empirical data suggests that, in fact, we are quite skilled at locating emotional feelings in different body regions. One striking study, testing across a multicultural (Finnish, Swedish, and Taiwanese) population of some 701 test subjects, found that subjects could reliably point to different body regions whose activity levels vary under different emotional circumstances (Nummenmaa et al., 2014). Subjects were provided with silhouettes of bodies and asked to color the regions where activity would expectedly increase or decrease were the silhouetted persons presented with range of emotional stimuli. Anger-inducing stimuli, for example, sees greatly increased activity attributed to the face, upper torso (especially around the heart), and lower arms (especially the hands). Fear shows a similar spread, but with only modestly increased activity in the arms and more in the belly region. Happiness shows widespread increases in activity across the body, while depression shows decreased

activity throughout the body, moreso in the limbs. Unsurprisingly, changes in the face are represented in all emotional situations; interestingly, however, depression appears the only instance of a decrease of activation.

Assuming, as the results suggest, that this ability to localize bodily activity during emotion episodes is cross-cultural, a natural question is: is it innate? In a follow-up study, the same group of researchers asked a population of Finnish children (ages 6–17) to complete a near-identical task (Hietanen et al., 2016). Older children produced more discernable maps, resembling those of the adult population (Nummenmaa et al., 2014), between differing emotional circumstances than did younger children. However, even the younger subjects reliably produced bodily maps that showed different areas of activation depending on emotional circumstances. Happiness, for instance, is consistently associated with heightened activity across the body. And anger sees activity in the chest, though the attributed levels of activity appear to rise with subjects' ages. Of course this doesn't prove that we have innate associations between affective feelings and their bodily causes. However, the fact that very young children are able to produce such maps fairly consistently suggests that even early on, we learn to read our bodies when reporting our emotions.

I suggest, then, that bodily sensations are crucial cues from which our metacognitive systems conceptualize our emotion states. When I notice that my heart has started thumping, that I'm clenching my fists and gritting my teeth, that I'm contorting my eyes and mouth into a Clint Eastwood-esque expression and breathing deeply, I attach the label "anger" to my overall state and utter (or think) something akin to "I'm angry". When you feel your cheeks bunch up, your heart beat a little quicker, and your posture straighten as you are compelled to reach out and hug your loved one, you attach the label "joy" to your overall state and utter (or think) something akin to "I'm happy". The ways in which yours and my emotional experiences feel distinct from one another is due to where, across our bodies, activity is taking place—and how this translates to a point on the state space of core affect.

We thus report on our emotions by reading our own bodies: muscles, bloodflow, guts and all. The ancient Roman haruspices would divine knowledge by inspecting the entrails of chickens and goats. Reporting on our own emotions is a bit like introspective haruspicy, albeit without all the mess and spilled blood. Emotions remain actually embodied phenomena, even when we report on them using our conceptual repertoires.

Up to this point, this chapter has investigated what are the kinds of internal cues (i.e., first-person psychological phenomena and their physiological causes) we rely upon in metacognizing toward emotion self-reports. The next section determines what kinds of external resources are available for the construction of self-reports—where “external” in this sense means deriving from features distinct from the changes in core affect and pick-up of affective qualities that prompt the organism to generate a self-report.

4.3 Dynamic episodes, static reports

Our situations are ever-changing. Our eyes flitter around different points in the visual environment. Our ears pick up on minute changes in the ambient soundscape. Our interoceptive systems constantly monitor pressure in our bladders and bowels and bronchi. Blood sugar and hormone levels rise and fall, memories are retrieved and manipulated in the background of consciousness, our legs tingle and prompt us to shift position in our chairs every few moments. The state of the living body is never static or rigid.

Nor are the environments in which our bodies exist. Other people enter and leave the room, phones buzz, cars pass by outside, birds fly into windows and dogs bark maniacally at them. Even if we could hold the body fixed in one position, geographically and physiologically, the range of stimuli to which our bodies react would almost never be the same between two points in time. And if you did manage to still your environment, good luck trying to still your mind: as testimony from insomniacs and users of sensory isolation chambers will

tell you, utter silence can be the most salient of stimuli.

Emotional experiences function to relate the ever-changing body to the ever-changing environment. It follows, then, that they too are dynamic by their nature. They extend throughout time, as do any other mental phenomena, but they also morph into one another. An emotional experience with one kind of affective character can become a very different one within a single episode. A toy example (not intended as physiologically precise) can help. Imagine a loud noise startles you. This raises your heart rate, tenses your muscles, and causes your hormonal system to release a dose of adrenaline into your bloodstream. Though it may take several seconds for the adrenaline to start influencing your neural activity, the initial increase in activity throughout your muscular and circulatory system is enough to prompt you to form an emotion self-report. You deem yourself frightened. A few seconds later, when that adrenaline goes to work, you undergo another change in core affect—but, crucially, only a slight change, subtle enough that you aren't prompted to reassess your prior judgment regarding what emotion you're undergoing.

We can reproduce this episode in the analytic philosopher's preferred itemized format:

At time T0: Some affective quality is picked up during sensory registration.

At time T1: A range of bodily changes occur (heart rate increase, muscle tension, and release of adrenaline into the bloodstream); these bodily changes trigger a sharp change in core affect.

At time T2: You generate an emotion self-report whereby you deem yourself frightened.

At time T3: The adrenaline penetrates the blood-brain barrier and manipulates the function of your neurons, thereby causing a subtle change in core affect.

Now recall the Jamesian *constitution* claim, again from Chapter 2, section 2.2.1: emotions are constituted by felt bodily changes. When paired with the *specificity* claim (a particular set of

bodily changes specifies the phenomenality of a specific emotional experience), it leads us to acknowledge that since qualitatively distinct emotional experiences occur either side of T2, the emotion at T1 is numerically distinct from the emotion at T3. If an updated emotion self-report were generated at T4, it might be narratively identical to that generated at T2:

At time T4: You generate an emotion self-report whereby you deem yourself frightened.

But this second, newer report would refer to an affective state that is qualitatively distinct from that referred to by the report at T2. Nonetheless, the entire episode spanning from T0 to T4 is packaged up by the narrative self under the label *fear*. This seems to me a fairly normal scenario. When generating emotion self-reports, we tend to label a whole episode using a single emotion term. While your feeling-self experiences subtle affective changes such as those between T1 and T3 above, metacognitively speaking your narrative self does not.

In fact, it often seems to us that our emotions maintain for long stretches an iron grip over our psyches—especially negatively valenced ones—such that we fall prey to their suggestions despite our better judgments. Emotions, after all, used to be called *passions of the soul* in large part because they seem to happen not just in us, but *to* us (e.g., Descartes, 1985; Malebranche, 1997). While the world goes on around us, we might remain affectively frozen. Again, Terfry’s lyrics help to illustrate this: he speaks of being “strangled by guilt” and “caught in the headlights”. Without doubt, Terfry’s emotional experiences changed, if only minutely, throughout this episode. But he may not have noticed the changes, at least not coherently enough to form updated reports of them. (And can we blame him? His mind should have been elsewhere in those moments.) So, what the lyrics of this song describe for us is not simply Terfry’s emotional experience, but his emotion episode; one that’s particularly salient compared to other episodes of his life.

Why is it that, despite variances in emotional experiences within a given episode, we nonetheless label that episode using a single, encompassing concept? Why does one instance

of anger or joy seem to endure throughout minutes, hours, or longer still? More pressingly: what causes us to identify a temporal period *as* a discrete emotion episode? After all, we are constantly having emotional experiences. So how is it that we come to draw boundaries between one emotional episode and the next, or between “emotional” and “unemotional” moments?

The next two subsections each present a candidate answer to these questions. Firstly, in the spirit of the cognition-arousal theory, our range of available emotion concepts shapes how we conceptualize and report on our emotional experiences and the episodes in which they are most salient (section 4.3.1). Secondly, we typically associate the presence (to consciousness) of certain objects of emotion with the onset and endurance of emotion episodes (section 4.3.2). One or both of these components can determine the beginning and end points of a given emotion episode insofar as the narrative self reports on it.

4.3.1 Emotion concepts

Before diving into a discussion of emotion concepts, let me reiterate an issue raised in Chapter 2, section 2.2.3: philosophy and the cognitive sciences have not yet landed on solid ground regarding the nature of concepts. This fact partly undercuts the utility of Schachter and Singer’s cognition-arousal theory. It also is the largest obstacle to understanding precisely how we label our emotion states and whether, in fact, our emotion self-reports ever fully capture the emotional phenomena they purport to describe.

The words we use everyday to label and describe our emotion episodes, our emotional experiences, and our emotional reactions cannot be listed easily. Investigations into the range of terms used by English speakers to name particular emotional phenomena—not to distinguish here between episodes, experiences, and reactions—haven’t agreed on any count more specific than somewhere between 100–250 (see Cowie & Cornelius (2003, 12-13) for a round-up of the numbers). But even shrinking the set of emotion categories by cluster-

ing terms together in a fairly heavy-handed manner won't see those categories collectively encompass the range of emotional phenomena we experience:

The problem is exacerbated by the fact that despite their number, everyday categories do not capture every shade of emotion that people can distinguish. A vocabulary of 60 words or so may be too large to be tractable, but it is too small to capture distinctions that human observers make, and feel are important. Pictorial art provides a neat way of making this point. Artists revel in expressions that convey an emotional state which is very easy to identify with, and yet very hard to verbalise. (Cowie & Cornelius, 2003, 13)

Numbers of emotion terms available to emoters also vary between languages. As we've seen, numerous emotion terms do not translate coherently between languages or cultures. It is highly likely that a great many terms we readily apply to our emotion episodes depend on the language(s) we speak.

Before offloading this issue to the local linguistics department, it's worth our exploring an idea that promises to link etymologically emotion concepts to emotions proper. It may be that at least some emotion concepts are *embodied* in nature. Paula Niedenthal and colleagues have recently collated evidence from psychological and neuroscience studies that speaks in favor of a theory of emotion concepts as embodied simulations: "the activation in the body's sensorimotor and affective systems in many cases constitutes the conceptual content itself" (Niedenthal et al., 2014, 242). The idea here is that just as a qualitative instance of joy is due to activation of particular neural systems pertaining to sensorimotor and affective activity, those neural systems also embody a representation of joy when activated in unison.³⁶ In effect, when you think of joy, the very same neural circuitry responsible for its embodied experience comes online and is co-opted by your cognitive systems. Embodied emotion

³⁶Antonio Damasio offers a precursor to this theory that he calls the neural "as-if loop": a neural system's activation mimics the activation of a sensorimotor system, leading the organism to feel as if its body is undergoing some activity (Damasio, 1994, 1999; Bechara et al., 2000).

concepts, then, are distributed throughout modality-specific regions in the brain.

The evidence Niedenthal and colleagues appeals to derives from studies on perceiving and interpreting the emotions of other people. It suggests that in attempting to understand another person's occurrent emotion state, we mimic, internally, her observable facial and bodily positioning. For instance, when activation of the somatosensory cortices is inhibited via transcranial magnetic stimulation (TMS), test subjects show diminished accuracy when asked to identify facial expressions. Similarly, subjects are better at identifying an emotion when perceiving a facial expression that is concordant with bodily posturing (e.g., an angry-looking expression with an angry-looking body) than when identifying an emotion in someone whose expression and posturing are discordant (e.g., a happy-looking expression on an angry-looking body). If emotion concepts are embodied in the way suggested above, then this evidence suggests that emotions are perceived holistically, using information from throughout the body (Niedenthal et al., 2014, 245-246; see also Aviezer et al., 2008).

Note, too, that this conclusion concords with that proposed by Nummenmaa et al. (2014) and Hietanen et al. (2016), discussed above in subsection 4.2.2. In fact, we can draw a plausible inference from the pairing of these two conclusions. By virtue of our ability to trace our emotional experiences through our emotional reactions (the bodily maps view), we can learn to make reverse inferences regarding the emotional experiences of others by perceiving their overt emotional reactions and simulating them by activating our own sensorimotor representations of our bodies. Perceiving the slouched shoulders, drooping cheeks and lethargic gestures of a colleague, we simulate how those posturings would feel for ourselves. That simulation arms us with a virtual feeling that we associate with the term *sadness*. Noting this, we ultimately determine that our colleague is sad. Empathy, then, is a natural mechanism that attunes our perceptual system to the task of emotional mindreading.³⁷

Skeptics will be quick to point out that identifying another person's emotion state often

³⁷Malebranche and Hume reached very similar conclusions regarding what they called *compassion* (Taylor, 2013).

doesn't bring with it a corresponding emotion proper. I can see a sad person on the street, indeed see *that she is sad*, yet remain perfectly chipper and upbeat in my stride. On the flip side, when I'm particularly irked about something, nothing infuriates me more than seeing others express joy. But the view just offered appears to collapse *perception of emotion* into *emotional contagion*. Luckily, we can avoid this by noting the difference between the emotional experiences borne of actual versus virtual emotional reactions. Niedenthal et al.'s theory infers that sensorimotor activity is *simulated* during perception of emotion, not enacted. The body proper does not undergo wholesale changes—or, in James's terms, the simulation process does not involve a “diffusive wave” of bodily activity. Therefore, although the simulated emotional experience is real, it is attenuated in comparison to the perceiver's present, overriding emotional experience. When it comes to assessing our own emotion states, the more salient evidence will come from the dominant emotional reaction triggered by our actual bodily state.

If indeed embodied emotion concepts are grounded in our own (actual or simulated) emotional experiences, then our emotion words *mean* those experiences. Once again, then, this suggests that we appeal to our own feelings as primary components of emotion episodes. Emotion concepts pick out emotions either by our attending to actual emotional experiences or simulating them (if faintly) through modality-specific neural systems.

4.3.1.1 Discrete concepts for gradient phenomena

Niedenthal et al.'s theory of embodied emotion concepts is very much a live option. Unfortunately, it threatens to cause further headaches for us in our quest to explain why we appeal to discrete emotion concepts to describe non-discrete events. Consider the logical order when we embed it in a primitivist framework:

P1 The body is capable of undergoing indefinitely many different patterns of change.

- P2 Emotional reactions are grounded in bodily changes triggered during sensory pick-up of affective qualities (via interoceptive monitoring of exteroceptive activity).
- P3 Emotional experiences are grounded in emotional reactions.
- P4 Emotion concepts are grounded in emotional experiences (= feelings) and emotional reactions (= bodily changes).
- C We employ a restricted range of discrete emotion concepts to label our emotion episodes.

As this argument stands, the conclusion does not follow from the four premises. Since emotional experiences (indeed all affective states) blend into one another like hues on a color spectrum, there do not appear to be any principles by which to force divisions between kinds of emotional experiences. The concept *angry* is no more precise in its extension than is the concept *blue*: for either determinable term, there are indefinitely many determinates. So, if the premises and the conclusion are true—and I’ve labored to demonstrate their plausibility throughout this and the previous chapter—then we must be missing something between P4 and C. Some fifth premise must explain how we derive discrete emotion concepts from gradient emotional experiences. How do we find joints between concepts that don’t actually appear between the phenomena those concepts represent?

Here I can only offer very speculative suggestions towards an answer. The next chapter will offer more thoughts on this, too. For now it is helpful to turn to yet another extant theory. *Basic emotions theory*, also known as *affect program theory*, asserts that paradigm cases of those phenomena people generally refer to as “emotions” are best understood as coordinated patterns of bodily activity (both physiological and behavioral).³⁸ Proponents

³⁸There are roughly as many varieties of basic emotions/affect program theories as there are theorists proposing them. Some of the more impactful proponents include: Ekman (1980, 1992b, 1999); Ekman et al. (1972); Ekman & Friesen (1986); Griffiths (1990, 1997); Izard (1977); Scarantino & Griffiths (2011); Tomkins

of primitivism can think of these *basic emotions* or *affect programs* as large-scale emotional reactions: the kinds of “diffusive waves” of bodily activity that James (1884, 1890) talks about. Paul Griffiths offers a concise summary of these phenomena:

The central idea of affect program theory is that emotional responses [i.e., emotional reactions] are complex, coordinated, and automated. They are complex because they involve several elements. These are usually taken to include (a) expressive facial changes, (b) musculoskeletal responses such as flinching and orienting, (c) expressive vocal changes, (d) endocrine system changes and consequent changes in the level of hormones, and (e) autonomic nervous system changes. (Griffiths, 1997, 77)

It’s worth noting that Griffiths doesn’t commit basic emotions/affect program theory to any particular position regarding the relationship between physiological and affective phenomena. But in any event, its proponents hold that physiological and behavioral factors such as (a)–(e) provide a kind of Archimedean point for emotion research. Other features may be important to a given instance of emotion, but these are not the key features for tractability. Thus, as if an afterthought on behalf of basic emotion theorists, Griffiths continues: “Emotion feelings and cognitive phenomena such as the directing of attention are obvious candidates to be added to this list” (Griffiths, 1997, 77). When this theory’s proponents talk about emotion, they are referring to what I’ve been calling *emotional reactions*.³⁹

A key claim of basic emotions theory is that there exist a small number of affect programs that respond in distinct and fairly fixed manners to the ranges of stimuli to which they are attuned. Lay terms such as *anger*, *fear*, *joy*, *surprise*, and *disgust* purportedly track discrete affect program modules. (There are numerous lists of basic emotions in the literature. A

(1962, 1963). In the present discussion, I do not take care to distinguish between varieties; a surface-level description suffices. But for the most nuanced and updated iteration, see Scarantino (2014).

³⁹This diversion from the primitivist theory, which centralizes conscious experience in identifying emotion proper, mirrors a key point of departure between James’s and Lange’s theories (see Chapter 1, section 1.1.2.1).

conservative count is five basic emotions/affect programs, though some researchers identify fifteen or more.) In turn, each module is attuned to a different kind of ecological property: “The fear response is adapted to dangers, the disgust response to noxious stimuli, the anger response to challenges, the surprise response to novel stimuli” (Griffiths, 1997, 89). Some of that attuning comes built into the organism at birth, while other attuning occurs over the course of development, via associative mechanisms or cognitive operations. For example, while many primates display innate fear reactions to the presence of snakes (e.g., Le et al., 2013; Öhman, 2009; Shibasaki & Kawai, 2009), it takes humans a lot of learning to react with anxiety to the prospect of being fired. However they’re sensitized to stimuli, and regardless of which local events possess those ecological properties within a given environment, affect programs function to respond to events that are ecologically significant to the organism (Griffiths, 1997, 89).

If affect program modules exist, then there is a determinable range of emotional reactions that humans and other animals are prone to repeating more often than other kinds of emotional reactions. It is likely that the oft-felt emotional experiences triggered by a given affect program module on different occasions will resemble one another very closely. If so, then the individual will probably assign a label to that module’s output. And if all neurotypical humans have the same set of affect program modules, then we can empathize with one another’s affect program-triggered emotional reactions. Thus, we find comparable words across cultures for emotional experiences triggered by particular affect programs. In terms of embodied emotion concepts, these common experiences ground universal emotion concepts such as *anger* and *fear* and *joy*.

Basic emotions theory is obviously a limited resource for explaining emotion concepts. If, as cited above, English speakers employ somewhere between 100–250 words to describe their emotion episodes, then it’s unlikely that we’ll find a related affect program module for each of those words. The major plotline of Griffiths’s book distinguishes between affect pro-

gram modules and numerous other phenomena that laypersons call emotions. Although some theorists prefer to exile from the realm of emotionality anything that isn't explained in terms of affect program mechanisms, Griffiths argues that such phenomena "are no less worthy of investigation than those which fit the affect program model" (Griffiths, 1997, 101). I agree. However, Griffiths ultimately argues that the two distinct groups of phenomena to which we erroneously apply the same term (*emotion*)—affect program modules as "basic emotions", and higher cognitive "emotional" phenomena—do not necessarily share any features with one another. Therefore if *emotion* is to remain a genuine scientific term, it must apply to one group but not the other. Proponents of primitivism need not make such a distinction, since only emotional experiences, not emotional reactions (whatever their causes), count as emotions proper. And not all emotional experiences can be due to affect program modules. Thus, not every legitimate emotion is due to the activation of affect program modules.

Whether other emotional experiences, triggered via mechanisms other than affect program modules, ground embodied emotion concepts can be discovered through scientific investigation. One possibility is to modify the study design used by Nummenmaa et al. (2014) and Hietanen et al. (2016) in their bodily map studies to test for non-basic or gradient emotional experiences. These researchers take their findings to provide evidence in favor of discrete emotional feelings, as proponents of basic emotions theory would expect. However, their assumption of discreteness is built into their experimental design: subjects are asked to respond to a series of emotionally significant vignettes, each of which is shaped to convey a preconceptualized emotion. Since the experimental stimuli are restricted to a handful of varied situations, it is unsurprising that results would reflect the qualitative distances between the stimuli. Thus, while their data are consistent with basic emotions theory, they do not provide evidence against a gradient account of emotional experiences such as those proposed by James, Russell, or the primitivist theory.

Without first gathering the empirical data, it's difficult to tell a fuller story of how various emotion concepts relate to the experiences they function to convey or describe. We should also be open to the revelation that some commonly employed emotion terms do not genuinely track emotions proper. Perhaps it's time to send this problem on to the local linguistics department.

4.3.2 Objects of emotion

When it comes to our emotions, we have a strong tendency to offload the blame. If I'm angry at you, it's because—from my perspective—you were a jerk to me. If I'm happy with you, it's because you did something to benefit me. If I'm scared of you, it's because you're somehow threatening to me. *I* don't cause my own emotions, it seems; the rest of the world causes me to have them. And even if I concede that my thoughts cause my emotions, that's only because my thoughts are about the kinds of things in the world that matter to me. All this is to say that when we deem ourselves emotional, we typically identify some *object of emotion*.

Before continuing, I want to remind you that our present focus is emotion episodes, not emotional experiences. Recall from Chapter 3 that emotional experiences contribute to perceptual processing involved in object perception. To say that emotional experiences have objects of emotion would invert this relationship. Rather, an emotional experience partly *determines* an object of emotion.

This means that the cause of an emotion episode (that thing that triggers its onset via interoception of a state set up by exteroception) should be distinguished from that episode's object of emotion. If a bear-shape crosses my retinae and triggers a fear reaction, the ensuing emotional experience leads to my perceiving a frightening bear. This frightening bear is then the object of emotion throughout my fearful emotion episode. Per Russell's terminology (see 4.2.1 above), in perceiving the object as a frightening bear, I attribute to it the affective quality of being frightening.

I contend that the conscious presence of an object of emotion contributes to the endurance of an emotion episode. Two conditions must be met here. Firstly, one's attention must stay fixed on the object of emotion. If my perceptual or cognitive systems cease to represent or track the object of emotion (say, if the frightening bear wanders off to a safe distance), then the object of emotion no longer contributes to the longevity of my emotion episode. Consider the sigh of relief that marks the end of an episode of fear. (Of course, this is not to say that one's bodily feeling cannot endure after the object of emotion departs. Since affective change is due to bodily change, if the bodily symptoms take some time to reside, the corresponding affective state can hang around a bit longer. Perhaps you've noticed your hands continue to shake even after you've acknowledged the end of an episode of fear.)

Secondly, if I re-perceive the triggering object or event such that it undergoes a radical change in attributed affect, this will mark the transition from one emotion episode to a qualitatively distinct episode. In essence what happens here is that the object being perceived (the bear *qua* perceiver-independent physical entity) is recategorized as a different object of emotion: it's no longer a frightening bear but, perhaps, an awesome bear, or a cuddly bear, or a nuisance bear. This new object of emotion exists as part of a new emotion episode that will be differently conceptualized during the generation of an emotion self-report.

This second claim is perhaps unorthodox within the theoretical literature. John Deigh gives an example intended to demonstrate that objects of emotion persist while their emotionally significant properties change:

What's going on in a dog's mind when he growls at someone? Suppose, for example, you need to enter your neighbor's yard, but just as you approach the gate, their dog growls at you. What excites the dog's growling is his perception of you as you are about to encroach on his territory. He senses something invasive about your behavior that he would not sense in someone he knows and has affection for. Your appearance in his perceptual field triggers this sensitiv-

ity, and as long the condition lasts so does the growling. In fact, you could be someone whom the dog knows and likes but initially does not recognize. In that case, upon recognizing you, the dog will immediately change his attitude. He will stop growling and relax. His back, which would have been straight and stiff, will slump, and he will begin to wag his tail. Throughout the episode, you are the object of the dog's attention, and the dog tracks you in the sense that his emotion is sustained or altered according as his perception of you remains steady or changes. And what remains steady or changes in his perception of you is his sense of the invasiveness of your behavior. That sense could become stronger as you encroach further on his territory, or it could disappear altogether as soon as he recognizes you. (Deigh, 2004, 20)

This contradicts what I've said above about distinguishing the frightening bear from the non-frightening bear. To retell this story from a primitivist perspective, we would say that the dog transitions from one emotion episode to another. In the first, he picks up on affective qualities associated with a threat to his immediate well-being (*viz.* a conspecific entering his territory). In the second, he picks up on affective qualities associated with benevolence and affection. Each of these two distinct episodes has a unique object of emotion depending on which affective properties the dog picks up on during his repeated acts of perception.

You may not yet be convinced that my construal is preferable to Deigh's. Indeed, in terms of objective things in the world, it is one and the same thing—you—that the dog sees first as invasive, then as friendly. But while this is a solid metaphysical fact, it isn't necessarily at work in the perceiver's experience. Recall that primitivism subscribes to the theory of affordances as perceptual content. An object is perceived as what it is by virtue of its ecological relationship to the perceiver, by what it *affords* the perceiver. So, if at T1 the dog perceives you as invasive, and at T2 he perceives you as friendly, then between those two times he perceives two distinct sets of affordances (*attack-ability*, perhaps, shifting to

be-patted-by-ability). And, if he perceives two distinct sets of affordances, then with respect to the contents of his perceptual states, he perceives two distinct objects.

But what if the dog's psychological states do concord with the metaphysical facts, namely that at both T1 and T2 he is perceiving one and the same entity? Adult humans are pretty good at object permanence, which is the ability to continue tracking an object when it is occluded or rendered otherwise imperceptible. Less extreme, yet still intriguing, is the closely related capacity to continuously identify some object X as X in the face of property changes. (This tendency is at the heart of Theseus's Paradox.) Since dogs too are apt at object permanence (e.g., Gagnon & Doré, 1994; Miller et al., 2009; Triana & Pasnak, 1981), it's a safe bet that they share the latter capacity with us and probably many other species. One could object, then, that since the neighbor's dog perceives you as the same entity at both T1 and T2, despite your radical change in affective qualities, we shouldn't think of the object of emotion as differing between the two episodes. Indeed, perhaps we should revert to Deigh's perspective and identify a singular episode, involving a change in the dog's emotional attitude toward a single object of emotion.

The primitivist can respond here by appealing again to the distinction between the physical and affordance contents of a given perceptual state. This distinction is similar to one Anthony Kenny (1963) draws between particular and formal objects. For Kenny, the formal object of an emotion is the property in virtue of which the emotion is what it is. The formal object of anger is an insult or offense; the formal object of fear is a threat to one's well-being; the formal object of joy is some benefit one receives or will receive. The particular object, on the other hand, is the physical thing to which the formal object is attributable. In Deigh's example, you are the particular object by way of which the dog perceives first invasiveness, then friendliness (where invasiveness and friendliness, or *attack-ability* and *be-patted-by-ability*, are the formal objects of the dog's emotions at T1 and T2 respectively). Subscribing to the perceptual theory of affordances invites the primitivist to speak in terms of formal ob-

jects according to their underlying affordances. Since affordances ground object perception, and affective qualities are species of affordances, then specific formal objects ontologically depend on their specific underlying affordances. Thus, per primitivism, the dog perceives you (the particular object) differently, in that he perceives you as affording him different interactions (different formal objects), between T1 and T2. If there's a different formal object at T2 than at T1, then these time stamps pick out different emotion episodes.

To sum up this section: I suggest that emotion concepts and objects of emotion both operate in the generation of emotion self-reports. Both phenomena prompt us to identify a period of time as an emotion episode, and both help to ensure that such episodes persist throughout an extended time period. However, their contributions lead to the narrative self's reports of emotions that do not necessarily match the feeling-self's experiences: while the events experienced by the feeling-self are dynamic, an emotion self-report describes an affectively static episode.

4.4 How to feel bad about yourself

There's a certain type of emotion episode that is prioritized by people interested in moral psychology, ethics, and related spheres. Emotions such as guilt, shame, and sorrow are interesting, and particularly tricky to explain, since they involve the self twice over—that is, they are *self-reflective*. Other, non-reflective emotions seem easier to understand, at least with respect to their intentionality. Jan Slaby articulates the intentional relationship of a non-reflective emotional experience:

While afraid, you experience something as dangerous and at the same time 'you' feel vulnerable in the relevant respect. But your experience of the danger is not separate from, but rather *consists in* your feeling thus vulnerable. Each emotional experience has that structure: Something affects you, and *thereby* you

feel affected by it. Your ‘minding’ and something else’s ‘mattering’ are constitutively interrelated – there cannot be the one without the other. (Slaby, 2008, 438, original emphasis)

Under a primitivist interpretation, the “you” in this example refers to the feeling-self. The emotional experience carries information bidirectionally, reflecting the state of affairs of the world insofar as it relates to the state of your embodied self. (Recall Perry’s notion of primitive self-knowledge from Chapter 2, section 2.2.4.)

An episode of guilt, or any other self-reflective emotion episode, doesn’t seem to fit Slaby’s description. The feeling-self is not the kind of thing that can stand as the object of emotion, since it is the very vessel through which emotions are experienced, and by way of which objects of emotion are apprehended. Indeed, the feeling-self cannot be divorced from emotional experiences. But clearly in cases of self-reflective emotions, we identify ourselves as objects of emotion. How can this occur?

Back in Chapter 2, section 2.2.4 I distinguished between the concept of self (or narrative self) and the feeling-self (or embodied self). I think this is our best way to understand self-reflective emotions: Although the feeling-self *has* the emotional experiences, the narrative self stands as the particular object of emotion.

Recall Terfry’s situation once again. He feels guilty for having cheated on his partner. He describes an emotion episode during which certain emotional experiences endure. He acknowledges that he has done something bad, something that violates the expectations of someone he cares about very deeply. By extension, he has violated his own expectations regarding how one should live one’s life (e.g., avoiding infidelity in monogamous relationships, or perhaps merely avoiding causing harm to others).

In cases such as Terfry’s, the emotional experiences will be informed by cognitive mechanisms, not sensory systems. The nature of this mental circuitry is the subject of ongoing scientific investigation. One possibility is that the mind’s associative mechanisms al-

low for socially or otherwise contingently constructed norms to become associated with the ecologically-attuned emotion mechanisms described in the previous chapter. (Basic emotion theorists and Jamesians sometimes attempt to describe emotional responses to moral stimuli in a similar fashion.) The key difference between self-reflective emotions and other social emotions is that the metacognitive system operates over one's concept of self, rather than one's representation of another person. Terfry's feeling-self feels the way it does because it picks up on certain affective qualities of his concept of self that are made salient by his metacognitive system.

I have not tried to provide a comprehensive account of the causal relationship between the concept of self and the feeling-self. I am not well placed to comment on this. However, it strikes me as obvious that just as the feeling-self can motivate behavior by way of its emotional experiences (and other affective and perceptual states), so too can the concept of self motivate behavior. In fact, this is probably a major benefit of the human capacity for narrative self-consciousness. Consider a tame example: I don't decide to chop onions because I'm hungry; I decide to chop them because I'm trying to cook a decent bolognaise sauce. In fact, I hate chopping onions. It makes me very emotional. So, I require my cognitive system, including my concept of self, to motivate me to act in ways that brings my feeling-self immediate displeasure for the sake of downstream gain. Thus, my cognitive systems function to override the behavioral prompts of my feeling-self, at least from time to time. More generally, being able to check my actions and deliberate over a coping strategy in the face of ecologically significant situations can lead to better outcomes than succumbing to the prompts of an emotional experience. In terms of the practical relationship between the feeling-self and the concept of self, West Coast rapper Ice Cube's motto fits the story: "You better check yo self before you wreck yo self".

Terfry's described situation is a bit like this, except it reflects his prior failure to check himself. He feels guilty because he did something he knows he shouldn't have, and his con-

cept of self is present to him as blameworthy. How he responds to himself depends on his set of moral norms, and these can be described in a separate project. But as for his emotion episode here, he has everything he needs to generate a self-report: negatively valenced, localized emotional experiences, a consciously salient object of emotion, and emotion concepts associated with numerous situated motifs and imagery. Such are the ingredients for self-reflective emotion self-reports. Such are what you'll require to feel bad about yourself.

Conclusion

This chapter has presented *emotion episodes* as psychologically complicated events, each having numerous affective, perceptual, and cognitive components. We humans have a striking capacity to report on our emotion states. To generate emotion self-reports, we rely on evidence from throughout an emotion episode. Such evidence includes emotional experiences as well as metacognitive testimony, maps of bodily feelings, stocks of emotion concepts, and objects of emotion. Furthermore, the act of generating an emotion self-report also sets the temporal boundaries of the corresponding emotion episode.

By now it should be very clear that emotion episodes are much more than emotional experiences. Recall, once again, the primitivist claim that emotional experiences are emotions proper. Yet the kinds of phenomena explained in this chapter are often referred to as emotions: guilt and pride and shame are placed in the same bucket of psychological phenomena as are anger and joy and sadness. The present account suggests that this lumping together is erroneous (cf. Griffiths, 1997; Izard, 2007). That is not to deny that self-reflective emotion episodes do not have identifiable emotional reactions and emotional experiences (though the nature of such experiences remains unexplained for now). The claim is merely that the phenomena picked out by common-language emotion terms are not homogeneous. Not all emotional phenomena can (or should) be described in the terms of cognition-arousal the-

ory. Lay understandings of *emotion* do not fully overlap their psychological counterparts. The following chapter will explore this in greater detail.

Chapter 5

Emotion is a natural kind, emotions are not

Introduction

In the sciences of the mind, connections between explananda, data, and explanantia are typically more objectively opaque than they are in the so-called hard sciences—physics, biology, and chemistry. Their explanations rely more heavily on inferences and theoretical posits. Sometimes these posits are tailored to the topic of research (e.g., emotion, perception, decision-making, or memory) rather than neatly derived from their disciplines' general commitments. This is because mind is largely impenetrable from outside of the first-person perspective. While current technology allows us to sequence genomes and manipulate molecules, even the best brain imaging tools of cognitive and affective neuroscience don't enable us to measure or identify a subject's emotional feelings or the intentional contents of her thoughts and perceptual states.⁴⁰ Despite a long history of attempts to bridge or

⁴⁰Nor does current brain imaging technology allow for truly precise spatiotemporal resolution of neural activity. While electroencephalography (EEG) offers good temporal resolution of changes in neural activity, its spatial localization is poor and requires heavy predictive interpretation. fMRI gets better spatial resolution, measuring voxels of about three millimeters, at the expense of temporal resolution (measured blood flow de-

bypass the gap between the physical and the mental, mind has proven shy and evasive of our probing methods.

As a science relying on inferences, psychology faces a number of foundational questions: What grounds any given inference from observable (e.g., neural, behavioral) states to directly unobservable (e.g., conscious) states? Or from unobservable to observable states? How do psychologists determine and justify the ways in which they classify their discipline's phenomena? What are the most basic elements of mind, or features of mindedness?

The previous chapters described emotion as a fundamental mental phenomenon: an affective state that relates the world as sensed to the organism that senses. It is a consequence of a general affect system's activity pairing with that of sensory systems. But by virtue of its unique functions within the domain of mental phenomena, an emotional experience does not reduce to a mere pairing of affect system activity and exteroception. Emotion is its own thing; it's one of the basic elements of mind. The present chapter provides an argument for emotion's status as a *natural kind* in psychology. Emotion names a category whose members all share certain characteristics, and whose non-members lack those characteristics.

Sometimes discussion of emotion as a natural kind is run together with a related question: Are discrete emotions natural kinds? Proponents of basic emotions theory often say they are: since *anger* and *fear* (and *joy* and *sadness* and *disgust*...) reactions are observable in and identified by members of all human cultures and many other species, each of these terms picks out a discrete kind. To back up behavior-based investigations by psychologists and anthropologists (e.g., asking people to classify an emotion based on a staged photo of a facial expression (Ekman & Friesen, 1971)), neuroscientists devote much of their time to exploring and imaging the neural underpinnings of these discrete categories. One hears stories of the

notes activity from roughly three seconds earlier). But in three millimeters of of gray matter in the human cerebral cortex there can be as many as 630,000 individual neurons! Furthermore, both EEG and fMRI lose precision when imaging subcortical regions. Meanwhile, invasive single-cell recordings can yield much more spatiotemporally localized results, but these results cannot be contextualized within larger neural pathways or structures within the individual subject.

amygdala as the headquarters of fear, the insula as responsible for disgust, and other sections of the mid-brain and so-called limbic system taking on roles across the different discrete categories of basic emotionality. If discrete neural systems are discovered to produce discrete emotional reactions, the thought goes, then those discrete emotional reactions individually constitute natural kinds.

These two questions are distinct and should be treated as such. While this chapter argues for emotion's status as a psychological natural kind, it argues against claims of the natural kind status of discrete emotions.

The chapter is split into three sections. Section 5.1 explores dominant views from the philosophy of science regarding what natural kinds are and how they differ from other kinds of kinds. One other kind of kind, *investigative kind*, is an important tool for philosophy of science and the practices of the sciences of the mind. It plays a different role than the metaphysical notion of natural kind. Each kind of kind boasts different criteria for membership. In this first section I argue that we can employ both concepts to understand and explore the natures of both emotion in general and discrete emotions.

With the conceptual field set up, section 5.2 builds the case that according to the primitivist theory, emotion is both an investigative kind and a natural kind. Its status as an investigative kind is secured by the fact that each instance of emotion can be explained by a common homeostatic mechanism, namely the cooperation of interoceptive and exteroceptive systems in the emotional organism. Moreover, emotion deserves natural kind status because each instance of an emotional experience shares a common essence, much as do natural kinds in disciplines, such as chemistry, that demand greater precision in distinguishing ontological units than does psychology.

Section 5.3 then turns to the nature, or rather natures, of discrete emotions. While certain individual emotion categories, specifically those explained by affect program modules, may qualify as investigative kinds, no individual emotion category constitutes a natural kind.

The reason for the latter claim, I argue, is that there is no nonarbitrary means of clustering emotional experiences into kinds based on their intrinsic features, namely their descriptions under a circumplex model of affect (introduced in Chapter 3).

5.1 How to find a natural kind

The purpose of identifying natural kinds is to understand how units of an ontology fit together and relate to one another. The idea that kinds are natural means that those categories group together in ways largely uninfluenced by human interests. To use Paul Griffiths's preferred example, within geology and chemistry *jade* is not a natural kind, because the term picks out two chemically distinct stones: jadeite or nephrite (Griffiths, 1997, 2004a,b). The practice of chemistry dictates that natural kinds are differentiated according to chemical constitutions of substances. In the case of the two kinds of jade, a chemical analysis of one does not give direct information about the other. The fact that humans tend to fashion jewelry and statuettes from either stone does not justify any treatment of them as alike in a scientifically tractable sense. Thus, while jade is not a natural kind, jadeite is one, and nephrite is one too.

Discovering natural kinds in chemistry is fairly straightforward. We have a spread of resources at hand to perform chemical analyses of rock samples, and we have fairly well-established principles for distinguishing an instance of one kind from an instance of another. Investigation directs us to two chemically distinct profiles of "jade", each of which will then prompt its own further chemical investigation. The split between jadeite and nephrite marks a joint in nature. There's little we can do to argue against the change in our understanding of the natural kind landscape—unless, that is, we decide that chemistry is subordinate to the science of jewelry making!

Discovering natural kinds in psychology is significantly more difficult. It is both a young

and an inference-based science. The principle phenomena it investigates are not *things* per se—not rock samples or molecules or cells—but states and events. When we think about mind, we think about what humans (and other animals) *do*: how we behave, how we learn and remember, how we perceive, how we think and feel. These can't be extracted from tissue or excavated from dig sites. They are observed from the third person and reported from the first person perspectives. Moreover, even within psychology they can be ostensibly described at different levels: ecological, computational, and implementational (Griffiths, 1997; cf. Marr, 1982). This lack of direct objective contact with mental phenomena makes classifying them into kinds exceedingly difficult. It also calls into question the methods by which we should classify them, and what we intend by the term *natural kind* as it relates to psychology as a discipline.

5.1.1 Three approaches to natural kinds

Carl Craver (2009) compares three approaches to understanding natural kinds.⁴¹ We can think of these as constituting a spectrum of possible approaches.

At one pole, *conventionalism* holds that phenomena qualify as natural kinds if their concepts are fairly useful for explanation, prediction, or control. If researchers find it useful to treat a group of phenomena *as* a kind for the purposes of their scientific investigation, then that group constitutes a natural kind. As such, natural kinds derive their statuses from the disciplines in which they are employed. What qualifies as a natural kind in cognitive psychology, therefore, may not be a natural kind to neuroscientists or social psychologists.

At the other pole, *essentialism* claims that natural kinds pick out groups whose membership criteria are necessary. Again, chemistry provides some easy examples. *Water* is a natural kind per essentialism since for something to qualify as water, it must have the chemical composition H₂O (at least in our universe; let's leave Twin Earth out of this).⁴² A piece of metal

⁴¹In fact Craver derives these three approaches from Kornblith (1993).

⁴²Hendry (2006) explores the philosophical implications of some chemical caveats of this criterion.

qualifies as *gold* if and only if each of its nuclei has exactly 79 protons. In these two cases, chemical composition and atomic number capture their respective kinds' essences. Something is a member of natural kind X if it has the right essence; if it has a different essence then it is not a member of X.

Somewhere between the two poles is the less elegantly named *homeostatic property cluster* (HPC) view. On this approach, "a cluster of properties constitutes a natural kind when the co-occurrence of the properties in the cluster is explained by a homeostatic mechanism" (Craver, 2009, 576). Griffiths's affect program modules qualify since they pick out psychophysiological systems that are reliably triggered by a restricted range of stimuli and reliably trigger a coordinated pattern of behavioral and physiological responses (Griffiths, 1997). Under an affect program theory description, my dog's fear module responds to sudden loud noises (such as thunder or fireworks) by causing the poor fellow to shake uncontrollably, pant, try to hide behind the toilet, and urinate indiscriminately around the house.⁴³ Veterinarians and dog owners will attest that such responses are not uncommon among dogs. As a child I had another dog who acted similarly, though she was too big to hide behind the toilet so she'd crawl under the bed or try to burrow under the fence in the yard. On the HPC view, the psychophysiological system that triggers or constitutes my dogs' reactions to thunder and fireworks is a homeostatic mechanism: a mechanism that "explains the co-occurrence of the properties in the cluster" (Craver, 2009, 578).

The HPC view is more conservative than conventionalism since its purported natural kinds are sustained by mechanisms, and such mechanisms force real divisions upon the world (Craver, 2009, 577). For conventionalism to work properly, its advocates need to agree to principles and background theories that safeguard consilience between disciplines, especially neighboring pairs such as biology and psychology. If consilience fails, then interdisciplinary projects can't get off the ground, as explanations across disciplines will not

⁴³If you're planning to adopt a rescue dog, make sure you don't have carpeted floors.

obtain. The HPC view, meanwhile, can boast consilience as a basis for explaining homeostatic mechanisms: standard practices in the sciences of the mind see researchers uncovering neural mechanisms that help to explain psychological events. (The inverse relationship is also crucial: see Hatfield (2009b).) Discovering distinctions between neural operations that were once treated as alike or unified (through, say, lesion-based studies of stroke or other brain trauma sufferers) then leads to reinterpreting psychological events accordingly.⁴⁴ Without at least some consilience between neuroscience and psychology, something like the affect program theory would not be verifiable. And the fact that such correlations between mechanisms and events are described across different disciplines puts significant pressure on neighboring disciplines to conform to the lessons of each other's discoveries. Natural kind designations within disciplines arguably follow this trend even for a radically liberal conventionalist.

At the same time, the HPC view is more liberal than essentialism: its proponents reject necessary, definitive essences as criteria for kind membership for the reason that they are too inflexible when applied to the largely variable groups identified in sciences such as biology and psychology (Craver, 2009, 577; cf. Griffiths, 2004a, 905). For essentialism to triumph, it must find some compromise between restricting natural kinds to neatly systematizable phenomena (such as the units of physics and chemistry), on the one hand, and on the other accepting vague definitions for numerous essential criteria. While the former extreme would expect all sciences to discover criteria comparable to molecular structure or atomic number, the latter would invite fuzzy boundaries between essentialist natural kinds and thereby fail as a means of carving nature at its joints.

Although Craver (2009) and some others (e.g., Prinz, 2004; Scarantino, 2009) present the HPC view as an alternative to essentialism—rightly so, I believe—some theorists con-

⁴⁴See Goodale (1995); Goodale & Milner (1992); Milner & Goodale (1995) for a striking example regarding a neural split between two functionally distinct visual subsystems. Damasio (1994) also discusses the weird and wacky case of Phineas Gage, whose behavior and personality changed drastically after a metal rod blasted through part of his brain.

sider it a subspecies of essentialism (e.g., Barrett, 2006; MacLeod, 2013). Griffiths even points out that the original projects for which the HPC view was developed are exercises in “psychological essentialism” (Griffiths, 1997, 188). The idea is that a property cluster explained by a homeostatic mechanism defines the essence of a kind. But after offering his own practical definition of natural kinds, Griffiths says that “nothing hangs on the use of the term *essence*” (Griffiths, 1997, 189). I think that, if pragmatically rather than metaphysically, it is useful to separate the two views—the homeostatic property cluster view and essentialism, as described here—rather than subsuming one under the other (*essence*’s status notwithstanding). It helps to keep distance between the demands of either approach to identifying natural kinds. For supporting reasons other than the ones discussed above, see Craver (2009).

Pragmatically, then, the HPC view is more realist than conventionalism but more instrumentalist than essentialism. The world doesn’t carve up neatly or entirely without human interjection—since associations must be acknowledged between instances of homeostatic mechanism activity in order to infer the mechanism’s reliability and therefore its status as a natural kind-sustaining system—but, at the same time, we don’t get distinct, perhaps incompatible divisions of the world as we move between disciplines.

These three views all face their own conceptual issues, and none of them entirely overshadows the others (see Craver (2009) for analyses). The purpose of this section has not been to promote, say, the HPC view over conventionalism or essentialism. The issues each view faces will not be explored here. Instead, acknowledging these three approaches will aid in analyzing extant arguments regarding the natural kind statuses of emotion and discrete emotions. Before moving to the particulars in sections 5.2 and 5.3, let’s look at how natural kinds function in contrast to other kinds of kinds.

5.1.2 Other kinds of kinds

Natural kinds are not the only categories useful for predicting, explaining, or controlling ourselves and our environments. Griffiths (2004a) also discusses *investigative kinds* and *normative kinds*.

He explains what he means by the term *natural kind*: “I use the term ‘natural kind’ to denote categories which admit of reliable extrapolation from samples of the category to the category as a whole. In other words, natural kinds are categories about which we can make inductive scientific discoveries” (Griffiths, 2004a, 903). In line with his HPC view, he claims that we need not clearly define such categories in order to subject them to scientific investigation. But, it appears, many philosophers consider essentialism (or at least strict realism) the default mode for natural kind categories. The term *natural kind* therefore “carries a lot of unwelcome baggage” (Griffiths, 2004a, 905). To distance his own construal of natural kinds from that of the essentialists, Griffiths proposes renaming them *investigative kinds*. He borrows this term from Ingo Brigandt, who writes:

An investigative-kind concept thus originates when a certain pattern among a class of objects is observed and it is assumed to be founded on some theoretically important, but yet unknown relevant mechanism that generates this pattern. An investigative-kind concept is associated with a search for the basis of this kind. A specific hypothesis about the nature of this basis might exist and motivate the introduction of an investigative kind concept and guide scientific research. A full theoretical account of the investigative kind can only be given after appropriate empirical study and might reveal a variety of complications. (Brigandt, 2003, 1309)

For something to be an investigative kind, then, it suffices that we observe repeated patterns of properties across instances of phenomena reliably enough to identify such phenomena

by those patterns. Subsequent investigation into the nature of the homeostatic mechanism underlying those property patterns will then inform us of whether we are exploring not one but multiple categories (e.g., discovering that jadeite and nephrite are not chemically identical), and whether we can derive other scientifically significant knowledge regarding the investigative kind.

This is Griffiths's response to critics who, he claims, have misunderstood his position regarding the natural kind status of emotion (see Griffiths, 1997). Griffiths is an eliminativist, arguing that the category *emotion* is analogous to the category *jade*: both provide partial references, picking out multiple groups of incommensurable phenomena. *Emotion* might pick out an affect program module or response, a cognitively sophisticated emotion schema, or some other psychological event. And of those more specific phenomena, it is questionable which—if any—are themselves natural kinds. Perhaps many instances of discrete affect program responses constitute natural kinds: my dog's fear-reaction to thunder seems pretty reliable and coordinated. But at least in instances of human emotion, the term *anger* won't pick out the same pattern of responses each time. On some occasions, one's emotional response will be an affect program response; on others it'll be acting out a social script with the same emotion name (Griffiths, 2004b; Izard, 2007). My "anger" response to your spilling coffee on me is very different to my "anger" response at the current presidential administration's policy towards asylum seekers or international trade or the environment. So, on Griffiths's theory, *anger* is not an investigative kind, as the term doesn't track a cluster of properties reliably explained by a homeostatic mechanism (such as a discrete psychophysiological system). The aim of identifying investigative kinds, he writes, "is to find categories that allow *reliable* predictions in a *large domain* of properties. The classic examples of natural kinds, chemical elements and biological species, meet these desiderata" (Griffiths, 2004a, 905, original emphasis). The variety of phenomena typically named "emotions" do not.

I see no reason why one cannot at once employ both essentialist natural kinds and va-

guer investigative kinds. Griffiths's conceptual division between the two terms allows both to coexist peacefully: natural kind discoveries are epistemic endpoints while investigative kinds play crucial roles in guiding research projects. Determination of natural kinds is a job for metaphysics, while investigative kinds are resources for philosophers of science (cf. Griffiths, 2004a, 906). When investigative kinds of certain phenomena lead to lawful criteria for inclusion in a category, they also lead to the discovery of natural kinds in the more realist sense. Think of the previous examples of water (essence: H₂O) and gold (essence: atomic number 79), natural kinds *par excellence*. Going forward, I will use the terms *natural kind* in the essentialist sense and *investigative kind* in the metaphysically weaker sense Griffiths proposes.

Natural kind talk also participates in a nuanced relationship with normative concerns. This is especially true of psychology: laypeople in the West believe that *emotion* is a meaningful word and act according to, and in response to, the phenomena they call emotions, as well as desires and beliefs. These all can be warranted or unwarranted according to circumstances. (My dog's fear reaction to thunder is unwarranted given his safe location indoors and surrounded by capable humans. My wife's belief that the Royal Wedding is worthy of her attention is self-evidently unwarranted.) Emotions are said to influence agency and decision-making; in an earlier chapter I suggested they influence pretty much all conscious events in human minds. Griffiths's critics, he writes, have shown a tendency to infer from his claim that *emotion* is not subject to a unified scientific explanation that he believes the term *emotion* is "unviable to any cognitive role whatsoever" (Griffiths, 2004a, 902). Yet diagnoses of mental disorders often point to malfunctioning emotion mechanisms as contributors to those disorders. Therefore, the term *emotion* carries significant normative weight; according to Griffiths's critics this arguably justifies its status as a natural kind term.

But eliminativism regarding a natural kind *term* does not at all imply that the disparate phenomena that term picks out disappear from one's understanding of the world! Nor does

a shared normative quality alone qualify membership in a general natural or investigative kind. So, Griffiths suggests, we can think of *normative kinds* as different means of categorizing phenomena than investigative kinds (Griffiths, 2004a, 908). Like investigative kinds, normative kinds can be open-ended categories; they guide normative projects rather than epistemic ones (Griffiths, 2004a, 908). *Human-caused climate change* may constitute a normative kind alongside an investigative kind. Evolving beliefs regarding the environmental impacts of human behaviors sees the term's intension and extension altered; it also motivates behavioral changes such as reducing carbon emissions and better managing waste and natural resources. But normative kinds need not meet the criteria of investigative or natural kinds in order to be objects of inquiry. A category can be a functioning normative kind in absence of an equivalent investigative kind. Perhaps specific scientific disciplines will have less to say about normative kinds than will humanistic disciplines, though that does not diminish their importance. Some theorists, such as Peter Goldie (2000) and Robert Solomon (1976, 2004), are interested in emotions primarily (or solely) as normative kinds.

The current project is descriptive rather than prescriptive. Although I am interested in showing that emotion is a natural (and investigative) kind, I am not interested—not here—in extending this to normative concerns such as revising moral standards for responding to one's own and others' emotions. It's worth following Griffiths in distinguishing investigative from normative kinds since it allows for a division of labor between descriptive and prescriptive projects. With that said, we can now turn to the question of whether emotion is a natural kind.

5.2 Emotion is a natural kind

The distinction made in the previous section between investigative kinds and natural kinds invites two distinct lines of inquiry into the status of emotion as a psychological category.

This section will demonstrate that emotion qualifies as both kinds of kind. However, the conception of investigative kind is employed by its major proponent, Paul Griffiths, in arguing that emotion is *not* an investigative kind. To safeguard my own argument that emotion is an investigative kind, it is prudent to first attend to Griffiths's eliminativist position on emotion's kind status (subsections 5.2.1–5.2.2). Positive arguments then follow for emotion's status as both an investigative kind and a natural kind (subsection 5.2.3).

5.2.1 Arguments against emotion's natural kind status

As mentioned previously, Griffiths is an eliminativist about emotion: the category, he says, does not pick out a single investigative kind, so it should be eliminated from the psychologist's vocabulary. He identifies three broad classes of phenomena that commonly cluster under the name "emotion": affect program responses, higher cognitive emotions, and social pretences (Griffiths, 1997, 14-16). (Carroll Izard (2007) similarly distinguishes between basic emotions and emotion schemas.) For any instance of an emotion that can be explained as an affect program response, there is another that can only be explained by appealing to social schemas and cognitively sophisticated interpretations of one's circumstances. If any of these phenomena qualify as investigative kinds, they are the categories of discrete affect program responses: affect program "fear"; affect program "anger"; and so on. But if they are lumped together with the other two classes of phenomena, then the category containing all three does not pick out a group of phenomena sharing a common homeostatic mechanism:

My central conclusion is that the general concept of emotion is unlikely to be a useful concept in psychological theory. It is meant to be a kind of psychological process that underlies a certain range of human behaviors. But there is no one kind of process that underlies enough of this behavior to be identified with emotion. ... [W]hat we know about these phenomena suggests that there is no rich collection of generalizations about this range of phenomena that dis-

tinguishes them from other psychological phenomena. They do not constitute a single object of knowledge. (Griffiths, 1997, 14)

The thought here is that there is nothing suitably homologous between affect program “anger” and socially constructed “anger” that unifies them. Presumably, then, certain “emotions” would be better examined by sociologists than by neuroscientists, while explanations of others could benefit from the input of biologists and comparative psychologists. And so, psychologists and philosophers should give up the term *emotion*, instead splitting their resources among projects of narrower scopes. While Griffiths (1997) claims that emotion is not a *natural* kind, note from the previous section that by that term he means *investigative* kind. There are no homologous homeostatic mechanisms by which to explain each phenomenon. In short, emotion is not a psychological investigative kind.

Amélie Rorty is another theorist who claims that emotion is not a natural kind (Rorty, 1980, 2004). She points out the disjointed history of philosophy of mind, with numerous divisions and reconceptualizations made between emotional phenomena according to trends and conventions, such as the transition from emotions-as-passions to emotions-as-actions, and the different intentional characters of discrete emotional attitudes. Explanations of emotions’ causes divide into physicalistic or intentionalistic, with neither vocabulary reducible to the other; nor does either approach completely explain a given instance of emotion. This lack of conceptual consistency implies that “emotions” do not share anything like a common essence. Moreover, Rorty says, explanations of emotional episodes can vary in method and scope, depending on their appropriateness or rationale under their circumstances. Therefore, when viewed as contributors to and objects of knowledge, emotions do not constitute a unified category. In short, emotion is not an epistemological investigative (or natural) kind.

Rorty’s position is more pessimistic and extreme than is Griffiths’s. Griffiths offers a general theory positing three distinct categories of phenomena laypeople call “emotion”, and he puts forward affect programs as investigative kinds due to their evolutionary history in hu-

mans and other species. Each of these non-investigative kind categories, and each of the affect program categories, will require its own analytical and investigative projects. One can be an affect program “fear” researcher without calling oneself an emotion theorist. Meanwhile, Rorty suggests doing away with theories of emotions entirely: every instance of an emotion must be analyzed within a larger narrative of the emoter’s life at present (Rorty, 2004; cf. Goldie, 2000). Moreover, for Rorty it seems that normative features dominate explanations: whether a given emotion is appropriate given its circumstances, and what makes it so. The two theorists’ explanatory ambitions therefore are very different. Nonetheless, their general claims are equivalent: the term *emotion* does not carve out unified domain of psychological phenomena. Following Jesse Prinz (2004, 79), we can call this the *disunity thesis*.

5.2.2 Attempts to defend the unity thesis

Griffiths’s and (to a lesser extent) Rorty’s positions have attracted a number of responses from philosophers and psychologists in defense of emotion’s investigative kind status. Some theorists argue that emotions are all unified in some way, and thereby constitute an investigative kind. Let’s look now at two attempts to defend the unity thesis.

5.2.2.1 A first step: affect program as an investigative kind

Louis Charland deems Griffiths’s view “simply too radical and premature” (Charland, 2002, 533) and offers a revision that get us part of the way towards a unity thesis. Recall that Griffiths takes certain affect programs to constitute natural kinds. Charland argues that at a minimum, Griffiths should embrace the investigative kind category of all affect programs (that is, the category containing affect program “fear”, affect program “anger”, and any others). He achieves this by drawing on Jaak Panksepp’s (1998) “neurally based definition of emotion” (Panksepp, 1998, 47; quoted in Charland, 2002, 519). I will spare you the full definition, but one of its components is the claim that the neural circuitry of basic emotions “must be able

to generate affective feelings” (Panksepp, 1998, 49).⁴⁵ Acknowledging this feature, Charland writes: “Although basic emotions are themselves defined by clusters of neurobiological and physiological properties, what makes them all cases of *emotion* is the fact that normally they also share most or all of these more abstract general defining properties” (Charland, 2002, 520). In line with the HPC view, Charland (and Panksepp) conceives of basic emotions as homeostatic property clusters. He then argues that Panksepp’s definition is compatible with Griffiths’s description of affect programs. Therefore, Griffiths has reason to posit *affect program* as an investigative kind (and not merely certain affect programs as investigative kinds).

However, even if Griffiths agrees that *affect program* constitutes an affective kind, he will not go so far as to identify *emotion* with *affect program*. That is, safeguarding affect program as an investigative kind does not bring with it the investigative kind emotion. Such a revision of the vernacular *emotion* would invite confusion with respect to levels at which phenomena are explained, such as the ecological versus the cladistic level (Griffiths, 1997, 230-231). For example, an ecological explanation of fear—all fear—might describe it as “response to perceived danger”. (Appraisalist theories of emotion often categorize them in such ways, positing benefit- and danger-detecting mechanisms responsible for discrete emotional responses.) But for Griffiths, across numerous instances of “fear” the mechanisms that explain those instances will vary beyond what a cladistic account can unify (see Griffiths, 1997, 241-242); they will predominately involve either affect program modules or higher cognitive processes. Though a neural circuit may underlie affect program fear (such as fear of a looming predator), that same circuit will not, or not fully, underlie socially sophisticated fear (such as fear of a crashing economy). Although all affect programs may be of a kind (per Charland and Panksepp), not all “emotions” can be. The higher cognitive emotions require explanations of a different form.

⁴⁵Izard similarly stresses the “evolved feeling capacity” of basic emotions (Izard, 2007, 261).

5.2.2.2 A second step: irruptive motivation as an investigative kind

Prinz (2004) questions whether the categorical split between affect program responses and higher cognitive emotions here is warranted. He offers an illustration of Griffiths's position:

Griffiths regards the cognitive prerequisites of higher cognitive emotions as a tipoff. Unlike affect programs, higher cognitive emotions do not seem to be modular. Not only can beliefs influence a state of shame but they also seem to be obligatory for shame. If you do not believe you did anything wrong, you will not feel ashamed. Shame can be caused by beliefs and cured by beliefs. If you discover that your actions were beneficial rather than harmful, you can trade shame in for pride. (Prinz, 2004, 83)

According to Griffiths's theory, the major point of difference between some affect program response and some higher cognitive emotion like shame is a lack of shared homeostatic mechanism. Affect program responses are quickly triggered and operating, and for the most part they are cognitively impenetrable (Griffiths, 1997). Higher cognitive emotions such as shame are gradually triggered (following cognitive interpretation of one's recent actions), can unfold over a long period of time, and are sensitive to other cognitive events, as Prinz's illustration shows. While Griffiths says that both phenomena are examples of *irruptive motivations*, he strongly rejects the idea that this category names a scientifically useful investigative kind (Griffiths, 1997, 16). However, Prinz argues on functional grounds that irruptive motivation does constitute an investigative kind:

The very fact that affect programs and higher cognitive emotions can both be described as irruptive motivations *constitutes* a unifying causal mechanism. Irruptive motivation is a causal role that is responsible for correlations between many of the "superficial" properties of emotions. It explains why emotions seem passive, drive action, and influence practical reasoning in seemingly irrational

ways. (Prinz, 2004, 85)

Prinz's view is that a shared function is sufficient to unify two categories into one, namely the investigative kind irruptive motivation. And the category of irruptive motivation may indeed be the investigative kind category emotion.

Unfortunately, Prinz recognizes that this argument is "not fully satisfying" and he suspects it "will fail in the end" (Prinz, 2004, 85). Indeed, Griffiths (2004b) responds to this line of reasoning by appealing to consilience between psychology and biology with respect to how investigative kinds are identified. In biology, two fundamental sets of categories are *homologies* and *analogies*. A homologue is something that appears in two instances (say, members of two distinct species) that has descended from a common ancestral form. Manifestations can differ in form and function between instances, but they collectively constitute an investigative kind by virtue of their shared evolutionary history. A bat's wings and human arms are homologous in this sense: both shared an earlier mammalian forelimb as their common ancestral form. An analogue, on the other hand, is something that has the same function as, but different evolutionary history than, something else. A bat's wings and a bee's wings both enable flight, but they do not share a common ancestral form.

Griffiths's view is that explanations using biological investigative kinds show priority for homologies over analogies. Prinz's argument for the unified investigative kind irruptive motivation appeals to analogies (shared functions) rather than homologies (common homeostatic mechanisms). But this won't work if consilience between biology and psychology is to be maintained:

Any psychobiological theory of emotions *in general*... would have to be a theory of psychological analogies—traits that fulfill the same functions in relation to the environment. The categories that would be generated by such a theory, although they might enter into useful ecological generalizations, would be systematically unsuited to the distinctive purposes of psychology and neuro-

science. They would support induction and explanation of the wrong domain of properties... (Griffiths, 2004b, 237, original emphasis)

A conventionalist might respond by rejecting Griffiths's claim regarding the necessity of maintaining psychological–biological consilience. After all, there is a long-standing trend in psychology of explaining mental states in terms of functions. Functionalists such as Jerry Fodor (1975, 1980, 1983) stress the disciplinary autonomy of psychology with respect to both its explananda and its methodologies. Many functionalists therefore will see explanation by analogy as the default for their discipline, as Prinz has done here. Griffiths's view, meanwhile, looks more like a form of physicalism: the assumption that consilience between the sciences (if not explanatory reduction) demands that the special sciences follow the lead and logical structure of the hard sciences. Rather than developing such a response, however, let's grant Griffiths his point and allow the constraint that investigative kinds in psychology be determined according to biological protocols. That is, for any psychological phenomena to group together as an investigative kind, they must share a common ancestral form. Their homeostatic mechanisms must be homologous. With this constraint in place, what is the investigative kind status of emotions under primitivism?

5.2.3 Emotion's status under primitivism

Recall that primitivist emotions are affective states occurring when the general affective system responds to activation of exteroceptive systems. Emotions proper are not patterns of behavior or physiological changes; nor are they merely affects. They are feelings that carry information about the self–world relationship insofar as one's homeostasis is being or may be impacted. What is the kind status of these psychological events?

5.2.3.1 Emotion is an investigative kind

Under the primitivist theory, emotion is an investigative kind. We can see how it qualifies by applying Griffiths's and Craver's criteria from above. For something to be an investigative kind, its members must be explained according to a common homeostatic mechanism. Chapter 3 argued that all primitivist emotions are produced by a common homeostatic mechanism: interoceptive monitoring of exteroceptive activity. Coordination between these two systems (interoception and exteroception) leads to specific affective states that together constitute a unique category within the realm of affective phenomena. Since all instances of primitivist emotion are explained by way of this shared homeostatic mechanism, all instances of primitivist emotion group together as an investigative kind. Although other affective states exist (such as hunger, thirst, and bodily pains), these do not qualify as members of the investigative kind *emotion*, because they are not caused by activity of exteroceptors.

Some may take issue with this portrayal. As the primitivist theory is novel there are no direct responses available in the literature. However, a near-counterargument can be derived from Andrea Scarantino's (2009) analysis of core affect's natural kind status (responding to arguments from Barrett (2006) and Russell (1980, 2003)). He argues that *core affect* does not constitute an investigative kind because the ranges of phenomena the category describes are significantly heterogeneous compared to one another:

There exist no frequently co-occurring inductively and explanatorily important properties that instances of core affect tend to share by virtue of causal homeostatic mechanisms. This is not to say that instances of core affect do not share any properties. For example, all instances of core affect can be characterized as having some degree of valence and having some degree of arousal. The point is that no inductions or explanations of interest to scientific psychologists are licensed by this mere fact. (Scarantino, 2009, 953)

Scarantino here subscribes to the same HPC view Griffiths endorses, and the one according to which I've argued that primitivist emotion constitutes an investigative kind. But the category of emotion (per primitivism) circumscribes a slightly narrower range of phenomena whose descriptions derive from the core affect model (see Chapter 3, section 3.1.4). Perhaps those who agree with Scarantino's claim that the category of core affect collects too many causally heterogeneous members to constitute an investigative kind will also declare that the category of primitivist emotion similarly collects causally heterogeneous members.

An observation can bolster this argument against primitivist emotion's investigative kind status. Different exteroceptive systems contribute to the processes underlying emotion. Thus, one could assume that different kinds of causal mechanisms explain different kinds of emotions, namely modally specific ones (visual, auditory, tactile, etc.). If so, then strictly speaking an instance of "fear" generated through visual stimulation is causally distinct from an instance of "fear" generated through auditory stimulation. As an example, my dog's felt response when hearing a sudden boom of thunder (absent his seeing lightning) is of a different investigative kind than his felt response when seeing a large, looming predator (absent hearing a growl or similar). Since either instance of emotion has a distinct mode of exteroceptive activity (audition vs. vision), these instances do not share a common homeostatic mechanism. The same goes for comparing emotions across species.

But this response is pedantic. Just as biological investigative kinds must allow for variation among their members, so too must psychological investigative kinds. Moreover, membership in a kind does not rule out membership in a subsumed kind. *Mammal* is an investigative kind; so too is *human*, and *dog*. I'm not a dog; my dog is not a human; but we're both mammals. With respect to emotions, differences in exteroceptive specifics may lead researchers to posit distinct investigative kinds based on modality—similar ideas will be discussed shortly—but these divisions do not undermine the value of emotion as a broader investigative kind. Homeostatic mechanisms can be explained at different levels of abstrac-

tion to allow for both broad and narrow categories of investigative kinds.

Perhaps Scarantino is also a little too quick in claiming that the category of core affect does not license inductions or explanations of interest to psychology. There are likely important discoveries to be made in comparative psychology by employing core affect as an investigative kind. Even more so primitivist emotion. Employing the latter as an investigative kind can help to discover the ranges of environmental features to which different species are perceptually sensitive. Such inquiries have driven psychological and philosophical trends at least since William James (e.g., Gestalt psychology, ecological psychology, and contemporary embodied cognition). Empirical investigations can allow us to better understand how different creatures navigate and comprehend their *umwelts* or “subjective worlds” (von Uexküll, 1926). And anthropocentric sciences can benefit from these discoveries when it comes to, say, using animal models to understand human neuropsychology, as well as in refining inferences (and moral concerns) during pharmacological experimentation on animals.

To summarize the present argument: According to primitivism, all instances of emotion proper are explained by appeal to a common homeostatic mechanism (that is, they are homologous, not analogous, with each other). Moreover, the category of emotion plays an important role in producing inductions and explanations of interest to psychologists. Therefore, according to the criteria of Griffiths’s and Scarantino’s HPC view, emotion is an investigative kind.

5.2.3.2 Emotion is a natural kind

Perhaps emotion enjoys a more privileged position in the metaphysics of mind. I now want to make the case for emotion’s status as a natural kind, not just an investigative kind. The primitivist theory can allow for an essentialist definition of emotion roughly analogous to the essentialist definitions that make water and gold natural kinds in chemistry.

Again, for a category to be a natural kind its members all must share a common essence. According to the primitivist theory, all emotions proper are affective states. Moreover, each instance of emotion has a particular intentional character: it carries information about the self–world relationship. These criteria determine emotion’s status as a natural kind.

If an affective state does not have this self–world directed intentional character, it is not an emotion. However, this claim is not accepted universally. Derek Denton and his colleagues note the difficulty of defining the term *emotion* before offering the following brief construal: “If the word ‘emotion’ was represented by a pyramid, then at the base are the primitive or primordial emotions, in the middle the distance receptor evoked emotions such as anger, hate, fear, love, and at the apex are the emotions such as are experienced with the aesthetic delight of great art or music” (Denton et al., 2009, 501).⁴⁶ Those “primordial emotions” are the conscious manifestations of purely interoceptive processes, rather than the collaboration between interoception and exteroception:

The primordial emotions include thirst, hunger for air, hunger for food, pain, hunger for specific minerals, sexual arousal and orgasm, sensations accompanying impediment of visceral function, as, for example, micturition or defecation, desire for sleep after severe deprivation, and avoidance of change of body core temperatures etc. . . . These subjective elements of instincts subserve the vegetative systems. They are genetically programmed and guard the physico-chemical constancy of the internal environment of the body—the homeostatic process. (Denton et al., 2009, 501)

These phenomena are undoubtedly important and interesting. They may constitute the full range of phenomena that could be experienced by our friend the solipsistic bivalve who lacks exteroceptors (see Chapter 3, section 3.1.3.2). But, *contra* Denton et al., they are not emotions proper, because they fail to relate their organism to its environment in a meaningful way.

⁴⁶Thanks, again, to Peter Godfrey-Smith for pointing me to this source.

If a mental state with a self–world directed intentional character does not involve a feeling of some quality, it is not an emotion. Note, however, that this does not demand that an emotional experience represent any object to the emoter. An emotion is not a form of perceptual state, nor is it an evaluative judgment regarding the contents of a perceptual state. Instead, as stressed in Chapter 3, an emotion contributes *to* perceptual processes by *affectively* importing information about the self–world relationship. While advocates of the primitivist theory should hesitate to allow for any conscious mental state unaccompanied by an emotional experience in human minds, we can nonetheless grant that mental states lacking affective qualities are in principle identifiable. Perhaps, following Magda Arnold and other appraisalists, one can allow for unemotional states such as a “cold perception” of the world (Arnold, 1960a, 107). And history shows that functionalist cognitive science thoroughly struggles to find place for feelings in an information-centric ontology (e.g., Lewis, 1980; Nagel, 1974). One might form the belief that *I am in danger* without feeling fear. Few theorists would strive to contradict the claim that such states are not emotions proper, though they may offer differing justifications as to why they are disqualified.

Overall, then, I believe we have good reason to accept that emotion is a natural kind category. It is a fundamental unit in an ontology of the mental. It is not merely a subspecies of affect: its members distinguish themselves through their unique intentional character. This sees emotion satisfy the essentialist criteria for natural kind categories.

5.3 Emotions are not natural kinds

With the category of emotion now established as both a natural and an investigative kind, we can question whether there are narrower natural and investigative kind categories subsumed under emotion. It is probably easy to believe *prima facie* that numerous discrete emotions exist under a primitivist construal. After all, empirical evidence presented in the previous

chapter suggests that from a young age humans are adept at identifying differences in bodily feelings correlating with different emotional situations. Similarly, Izard (2007) collects citations from neuroscience and behavioral studies that together suggest an ability to discriminate between different basic-emotion feelings states that is analogous to (or emerges concurrently with) an ability to distinguish between distinct taste profiles (sweet, salt, sour, and bitter). Field and lab studies executed by proponents of basic emotions theory/affect program theory suggest that some limited range of emotions is available across cultures and species—an idea mentioned at many points throughout this and previous chapters. In short, at least some emotions, if only the basic emotions, often seem to stand out on their own. Being in a state of anger feels different than being in a state of fear or joy or sadness.

Yet concluding from all this that discrete emotions such as *fear* and *anger* and *joy* and *sadness* constitute their own natural kind categories is a little too quick. Moreover, most of the literature that makes or analyzes such claims fails to follow Griffiths in differentiating between natural and investigative kinds (via essentialism and the HPC view). This section maintains that distinction, and therefore questions separately whether discrete emotions constitute investigative kinds (subsection 5.3.1) and whether they constitute natural kinds (subsection (5.3.2)).

5.3.1 Some emotions may be investigative kinds

The primitivist theory of emotion describes emotions in terms of affective states. These states' affective qualities can be described according to James Russell's (1980; 2003) circumplex model of core affect (see Chapter 3, section 3.1.1). This is a dimensional model: an affective state can correlate with indefinitely many points along the model's axes. And since the specific affective quality is determined by changes in different parts of the body, the range of changes a given body can undergo sets the range of affective qualities that body can produce.

There are no discrete subregions of the affective circumplex. Emotion concepts pick out

approximate locations on the affective circumplex: vague designations whose boundaries are gradual rather than rigid. These locations represent combinations of valence and arousal values. What one calls a feeling of “anger” will likely be negatively valenced and have an arousal value above a “neutral” baseline. What one calls a feeling of “joy” will likely also have an arousal value above baseline, but will be positively valenced (i.e., a pleasant feeling). But neither “anger” nor “joy” picks out a rigid or bounded range of possible affective qualities.

Analogously, a color wheel sees hues blend into one another a little at a time. When asked to pick out “blue”, one can point to a general region, but one will have a tough time circumscribing a range that both includes all possible shades of blue and excludes all other hues such as purples and greens. In other words, there are no readily identifiable regions that specify discrete emotional feelings on the affect circumplex, just as there are no readily identifiable joints between “colors” on the color wheel. Emotions shade imperceptibly into one another just as do colors (cf. James, 1890, II: 448). Another way to put it is that emotion concepts are determinables, each of which has indefinitely many determinates. Some of those determinates may be shared between differing determinables, such that the emotion concepts we use overlap with one another at times.

One way of making sense of these emotion concepts’ lack of fixity is to propose that the terms do not name discrete emotions—they do not pick out clearly circumscribed regions on the affective circumplex—but, insofar as they track emotions, they constitute “emotion families” (Ekman, 1992a; Ekman & Friesen, 1975). Although instances of a basic emotion family may vary in certain respects (such as duration, severity, or number of expressive components), each member of an emotion family shares a reliably observable pattern of characteristics with other members (Ekman, 1992a, 172). But this does not solve the issue of classifying borderline emotional experiences because, as Ekman writes, the “shared characteristics within a family differ between emotion families, distinguishing one family from another” (Ekman, 1992a, 172). The family view of discrete emotions still demands space

between families. This renders it incompatible with the circumplex model of affect.

Another attempt to solve this issues sees basic emotions theorists sometimes treating ambiguous emotional responses as blends of basic emotions (see Ekman, 1992a). This allows for some cross-pollination between basic emotions theory and Russell's circumplex model. For example, Nelson Zagalo and colleagues used data from a survey-based study to locate the emotion term *disturbed* midway between *anger* and *fear* on Russell's circumplex, while maintaining that *anger* and *fear* are basic emotions per Ekman's theory (Zagalo et al., 2005). But this data derives from an experiment whose design assumed and sought discrete emotional responses from its subjects, who reported on their own emotion states. And as the previous chapter stressed, an emotion self-report is not an emotion proper. Therefore, Zagalo et al.'s attempt to bridge basic emotions theory with the circumplex theory of affect does not help to answer the question of whether there are any categorically discrete emotional experiences.

Proponents of discrete emotion theories (of which basic emotions theory is one prominent iteration) often speculate, or claim that they have discovered, that there exist discrete neural systems that are responsible for discrete human emotions. Conversely, many empirical studies produce either inconclusive data or data that is inconsistent across studies. Lisa Feldman Barrett offers a meta-analysis of literature from psychology and neuroscience and concludes that "it is difficult, if not impossible, to empirically identify the extensions of each emotion category" (Barrett, 2006, 45). In other words, emotion categories such as *anger* and *fear* and *joy* do not constitute investigative kinds—at least not on the basis of empirical evidence. She repeats this line in her recent book (Barrett, 2017), perhaps leading one to suspect that no groundbreaking, universally accepted discoveries of discrete neural systems have been made in the last decade.

So, should we now conclude that there are not any discrete emotion categories? No. Not yet, in any event. Firstly, lack of evidence in favor of a claim is not the same as evidence

against a claim. (If it was then the criminal justice system would operate a bit differently.) Cognitive and affective neuroscience are still young sciences—even younger than psychology—and we are still learning how to parse neural systems from one another, and how they contribute to mental events such as affective states. We can keep looking for discrete neural systems as our investigative techniques and technologies develop. In fact, *contra* Barrett, recent alternative interpretations of empirical literature see hope for discrete emotions theory (e.g., Griffiths, 2013; Izard, 2007; Scarantino, 2009; Scarantino & Griffiths, 2011).

Secondly, and more importantly as regards the primitivist theory of emotion, the idea that discrete affect program modules reliably explain patterns of behavior and physiological changes is fully compatible with the idea that a general affect system represents these patterns to consciousness. This obtains even if there are indefinitely many possible emotional experiences available to the organism. A furrowing of the brow, coupled with a tensing of the skeletal muscles and a quickening of the heartbeat, might contribute all the physiological signals required for the general affect system to produce an experience of anger. Repetition of this coordinated activity makes for repetition of affective feeling, which in turn motivates categorization. In short, affect program modules may ground embodied emotion concepts that are manifested as identifiable emotional experiences. This idea was explored in the previous chapter and is presented originally by Niedenthal et al. (2014).

Primitivism allows that certain emotions constitute investigative kinds. Certain emotional experiences may be reliably explained by common homeostatic mechanisms: discrete patterns of physiological change during or following interoceptive monitoring of exteroceptive activity. Note, however, that discreteness of homeostatic mechanism does *not* necessitate discreteness of emotional experience. The affective state that constitutes the emotion proper may not be circumscribed within the general affect system, even if its homeostatic mechanism is discrete according to a biological description of the organism. If we discover that a given category of emotional experience is explicable by a homeostatic mechanism (such as

an affect program module in the brain, or perhaps one distributed across the brain and the body) then that emotional experience category constitutes an investigative kind.

Treating certain emotions as investigative kinds is advantageous for understanding the circumstances under which emotions are or are not triggered, both in humans and other species. Imagine that we discover a discrete neurophysiological mechanism for producing fear experiences (e.g., LeDoux, 1996, 2002a, 2014). We can then investigate whether in fact this is a solitary mechanism or whether it can divide into yet more specific mechanisms. (Again, one investigative kind can subsume numerous others.) Perhaps we will then find modality-specific kinds of fear, or otherwise specific relations between emotional experiences and the various senses. For instance, Kyle Gagnon and colleagues investigated the influences of fear on different perceptual modalities. The title of their article is telling: “Fear influences perceived reaching to targets in audition, but not vision” (Gagnon et al., 2013). Other empirical investigations have led researchers to claim similar effects on perceptual modalities—especially vision—by emotion or affect (e.g., Barrett & Bar, 2009; Cole et al., 2012; Cola et al., 2013; Duncan & Barrett, 2007; Hills et al., 2011; Leder et al., 2011; Leibovich et al., 2016; Panichello et al., 2017; Todd et al., 2012; Zadra & Clore, 2011).

5.3.2 No emotions are natural kinds

If certain emotions may be investigative kinds, does that mean they may be natural kinds as well? No. In fact, regardless of whether any emotion turns out to be an investigative kind, its natural kind status does not hinge on its investigative kind status.

Recall that for a category to be a natural kind, all of its members must share an essence. And that essence must be distinguishable from the essences of non-members. (All members of the natural kind category of gold have atomic number 79; no member of that kind has a different atomic number.) So, do certain emotions boast essences unique to their categories? I don’t believe so. While members of investigative-kind emotions share common

homeostatic mechanisms, the criteria for natural kind membership are stricter. And recall from section 5.2.3.2 that all instances of emotion share an essence: all emotions proper are affective states with a self- and world-directed intentional character. Simply put, there are no other features that certain emotions enjoy while others do not.

We might attempt to divide emotions into “positive” and “negative” based on valence values, but then we would face issues regarding how to classify those emotional experiences that hover around the neutral point of the valence dimension. Likewise for divisions according to arousal values.

We might appeal to differences in the kinds of information different emotions convey to us. But epistemic divisions beg the question of discreteness to begin with. To assume that, say, fear is different from all other emotions *because* it informs us of potential dangers to our well-being undermines the project of discovering the roles emotions play in different circumstances. (Moreover, this would require explanation via analogy rather than homology. Although we haven’t discussed whether this is appropriate for identifying natural kinds, Griffiths’s argument from section 5.2.2.2 regarding investigative kinds gives us reasons for skepticism.)

James himself was unhappy with the idea of taxonomizing emotions based on their features, comparing it to the tedium of reading “verbal descriptions of the shapes of rocks on a New Hampshire farm” (James, 1890, II: ch. 25; Ellsworth, 2014). Taxonomizing natural kinds of emotions requires analyzing and comparing their features. But any differences in affective quality that we’d find would not mark out the boundaries required to group them into natural kinds. I conclude that individual emotional experiences do not constitute natural kinds.

Conclusion

This chapter has argued that the category of emotion constitutes a natural kind in psychology. Natural kinds are means of understanding the fundamental units in an ontology. Since psychology is a young science, identifying its natural kinds can aid in conceptualizing, designing, and interpreting the data from empirical investigations of psychological phenomena. If we recognize that emotion stands on its own, so to speak, among the mental events, we can avoid running into the kinds of challenges faced by theories that attempt to reduce emotions to, or otherwise describe them in terms of, other mental phenomena such as evaluative judgments or perceptual states or processes. Instead, emotion is its own thing and deserves its own investigative approaches.

However, this does not safeguard the natural kind status of individual emotions. Indeed, folk conceptions of anger and fear and whatnot deserve their own analyses, and examinations of at least some individual emotion categories can be informed by their statuses as investigative kinds. But that does not imply that these categories enjoy natural kind status. While emotion is a natural kind, emotions are not.

Bibliography

Adolphs, R. (2004). Emotional vision. *Nature Neuroscience*, 7(11), 1167–1168.

Al-Shawaf, L., Conroy-Beam, D., Asao, K., & Buss, D. M. (2016). Human Emotions: An Evolutionary Psychological Perspective. *Emotion Review*, 8(2), 173–186.

Arnal, L. H., Flinker, A., Kleinschmidt, A., Giraud, A.-L., & Poeppel, D. (2015). Human Screams Occupy a Privileged Niche in the Communication Soundscape. *Current Biology*, 25, 1–6.

Arnold, M. (1960a). *Emotion and Personality, Volume I: Psychological Aspects*. New York: Columbia University Press.

Arnold, M. (1960b). *Emotion and Personality, Volume II: Neurological and Physiological Aspects*. New York: Columbia University Press.

Averill, J. R. (1992). William James's Other Theory of Emotion. In M. E. Donnelly (Ed.), *Reinterpreting the Legacy of William James* (pp. 221–229). Washington, DC: American Psychological Association.

Aviezer, H., Hassin, R. R., Ryan, J., Grady, C., Susskind, J., Anderson, A., Moscovitch, M., & Bentin, S. (2008). Angry, Disgusted, or Afraid? Studies on the Malleability of Emotion Perception. *Psychological Science*, 19, 724–732.

- Barbalet, J. M. (1999). William James' Theory of Emotions: Filling in the Picture. *Journal for the Theory of Social Behaviour*, 29(3), 251–266.
- Barrett, L. F. (2006). Are Emotions Natural Kinds? *Perspectives on Psychological Science*, 1(1), 28–58.
- Barrett, L. F. (2017). *How Emotions Are Made: The Secret Life of the Brain*. New York, NY: Houghton Mifflin Harcourt.
- Barrett, L. F. & Bar, M. (2009). See it with feeling: affective predictions during object perception. *Philosophical Transactions of the Royal Society*, 364, 1325–1334.
- Barsalou, L. W. (1999). Perceptual Symbol Systems. *Behavioral and Brain Sciences*, 22, 577–609.
- Bechara, A., Damasio, H., & Damasio, A. R. (2000). Emotion, Decision Making and the Orbitofrontal Cortex. *Cerebral Cortex*, 10(3), 295–307.
- Bell, M. (2013). *Hard Feelings: The Moral Psychology of Contempt*. New York: Oxford University Press.
- Bower, T. G. R. (1974). The Evolution of Sensory Systems. In R. B. MacLeod & H. L. J. Pick (Eds.), *Perception: Essays in Honor of James J. Gibson* (pp. 141–152). London, UK: Cornell University Press.
- Brigandt, I. (2003). Species Pluralism Does Not Imply Species Eliminativism. *Philosophy of Science*, 70(5), 1305–1316.
- Buck 65 (2003). Tired Out. In *Talkin' Honky Blues*. Halifax, NS: WEA.
- Cannon, W. B. (1927). The James-Lange Theory of Emotions: A Critical Examination and an Alternative Theory. *The American Journal of Psychology*, 39(1/4), 106–124.

- Carey, S. (2009). *The Origin of Concepts*. Oxford, UK: Oxford University Press.
- Carruthers, P. (2009). How we know our own minds: The relationship between mindreading and metacognition. *Behavioral and Brain Sciences*, 32, 121–138.
- Carruthers, P. (2011). *The Opacity of Mind*. New York: Oxford University Press.
- Chalmers, D. J. (1995). Facing Up to the Problem of Consciousness. *Journal of Consciousness Studies*, 2(3), 200–219.
- Charland, L. C. (1997). Reconciling Cognitive and Perceptual Theories of Emotion: A Representational Proposal. *Philosophy of Science*.
- Charland, L. C. (2002). The Natural Kind Status of Emotion. *The British Journal for the Philosophy of Science*, 53(4), 511–537.
- Chemero, A. (2009). *Radical Embodied Cognitive Science*. Cambridge, MA: MIT Press.
- Clore, G. L. & Ortony, A. (2013). Psychological Construction in the OCC Model of Emotion. *Emotion Review*, 5(4), 335–343.
- Cola, S., Balcetis, E., & Dunning, D. (2013). Affective Signals of Threat Increase Perceived Proximity. *Psychological Science*, 24(1), 34–40.
- Cole, S., Balcetis, E., & Zhang, S. (2012). Visual Perception and Regulatory Conflict: Motivation and Physiology Influence Distance Perception. *Journal of Experimental Psychology: General*.
- Colombetti, G. (2014). *The Feeling Body: Affective Science Meets the Enactive Mind*. London: MIT Press.
- Cowie, R. & Cornelius, R. R. (2003). Describing the emotional states that are expressed in speech. *Speech Communication*, 40, 5–32.

- Craver, C. F. (2009). Mechanisms and natural kinds. *Philosophical Psychology*, 22(5), 575–594.
- Dagleish, T. (2004). The emotional brain. *Nature Reviews Neuroscience*, 5, 582–585.
- Damasio, A. & Carvalho, G. B. (2013). The nature of feelings: evolutionary and neurobiological origins. *Nature Reviews Neuroscience*, 14, 143–152.
- Damasio, A. R. (1994). *Descartes's Error: Emotion, Reason, and the Human Brain*. New York: Avon Books.
- Damasio, A. R. (1999). *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. New York: Harcourt Brace & Company.
- Darwin, C. (1872). *The Expression of the Emotions in Man and Animals*. London: John Murray.
- Davidson, R. J. (1994). Complexities in the Search for Emotion-Specific Physiology. In P. Ekman & R. J. Davidson (Eds.), *The Nature of Emotion* (pp. 237–242). Oxford: Oxford University Press.
- de Sousa, R. (2017). Emotion. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. URL = <http://plato.stanford.edu/archives/spr2014/entries/emotion/>.
- Deigh, J. (2001). Emotions: The Legacy of James and Freud. *International Journal of Psychoanalysis*, 82, 1247–1256.
- Deigh, J. (2004). Primitive Emotions. In R. C. Solomon (Ed.), *Thinking about Feeling: Contemporary Philosophers on Emotions* (pp. 9–27). New York, NY: Oxford University Press.
- Deigh, J. (2010). Concepts of Emotions in Modern Philosophy and Psychology. In P. Goldie (Ed.), *The Oxford Handbook of Philosophy of Emotion* (pp. 17–40). Oxford: Oxford University Press.

- Delabarre, E. B. (1943). A student's impressions of James in the late '80's. *Psychological Review*, 50(1), 125–127.
- Demos, E. V. (1992). Silvan Tomkin's Theory of Emotion. In M. E. Donnelly (Ed.), *Reinterpreting the Legacy of William James* (pp. 211–219). American Psychological Association.
- Denton, D. A., McKinley, M. J., Farrell, M., & Egan, G. F. (2009). The role of primordial emotions in the evolutionary origin of consciousness. *Consciousness and Cognition*, 18, 500–514.
- Deonna, J. A., Rodongo, R., & Teroni, F. (2012). *In Defense of Shame: The Faces of an Emotion*. Oxford, UK: Oxford University Press.
- Deonna, J. A. & Teroni, F. (2017). Getting Bodily Feelings Into Emotional Experience in the Right Way. *Emotion Review*, 9(1), 55–63.
- Descartes, R. (1637/1985). The Passions of the Soul. In J. Cottingham, R. Stoothoff, & D. Murdoch (Eds.), *The Philosophical Writings of Descartes*, volume I (pp. 325–404). New York, NY: Cambridge University Press.
- Dewey, J. (1894). The Theory of Emotion: (I.) Emotional Attitudes. *Psychological Review*, 1(6), 553–569.
- Doi, T. (1971). *Amae no Kouzou [The Anatomy of Dependence]*. Tokyo: Koubundou Ltd.
- Donnelly, M. E., Ed. (1992). *Reinterpreting the Legacy of William James*. Washington, DC: American Psychological Association.
- Duncan, S. & Barrett, L. F. (2007). The role of the amygdala in visual awareness. *TRENDS in Cognitive Sciences*, 11(5), 190–192.
- Ekman, P. (1980). *The Face of Man*. New York: Garland.

- Ekman, P. (1992a). An Argument for Basic Emotions. *Cognition and Emotion*, 6(3/4), 169–200.
- Ekman, P. (1992b). Are There Basic Emotions? *Psychological Review*, 99, 550–553.
- Ekman, P. (1999). Basic Emotions. In T. Dagleish & M. Power (Eds.), *Handbook of Cognition and Emotion*. New York: John Wiley & Sons.
- Ekman, P. (2003). *Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life*. New York, NY: St. Martin's Press.
- Ekman, P. & Friesen, W. (1975). *Unmasking the face: A guide to recognizing emotions from facial clues*. New Jersey: Prentice Hall.
- Ekman, P. & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, 17(2), 124–129.
- Ekman, P. & Friesen, W. V. (1986). A New Pan-Cultural Facial Expression of Emotion. *Motivation and Emotion*, 10, 159–168.
- Ekman, P., Friesen, W. V., & Ellsworth, P. C. (1972). *Emotion in the Human Face: Guidelines for Research and an Integration of Findings*. New York: Pergamon Press.
- Ekman, P., Friesen, W. V., & Simon, R. C. (1985). Is the Startle Reaction and Emotion? *Journal of Personality and Social Psychology*, 49(5), 1416–1426.
- Ellsworth, P. C. (1994). William James and Emotion: Is a Century of Fame Worth a Century of Misunderstanding? *Psychological Review*, 101, 222–229.
- Ellsworth, P. C. (2013). Appraisal Theory: Old and New Questions. *Emotion Review*, 5(2), 125–131.

- Ellsworth, P. C. (2014). Basic Emotions and the Rocks of New Hampshire. *Emotion Review*, 6, 21–26.
- Ellsworth, P. C. & Scherer, K. R. (2003). Appraisal Processes in Emotion. In R. J. Davidson, K. R. Scherer, & H. H. Goldsmith (Eds.), *Handbook of Affective Science* (pp. 572–595). New York, NY: Oxford University Press.
- Evans, D. (2002). The Search Hypothesis of Emotion. *British Journal for the Philosophy of Science*, 53, 497–509.
- Fehr, B. & Russell, J. A. (1984). Concept of Emotion Viewed From a Prototype Perspective. *Journal of Experimental Psychology: General*, 113(3), 4.
- Fellous, J.-M., Armony, J. L., & LeDoux, J. E. (2003). Emotional Circuits. In M. A. Arbib (Ed.), *The Handbook of Brain Theory and Neural Networks* (pp. 398–401). MIT Press, 2 edition.
- Fellous, J.-M. & LeDoux, J. E. (2005). Toward Basic Principles for Emotional Processing: What the Fearful Brain Tells the Robot. In J.-M. Fellous & M. A. Arbib (Eds.), *Who Needs Emotions? The Brain Meets the Robot* chapter 4, (pp. 79–115). Oxford University Press.
- Fodor, J. A. (1975). *The Language of Thought*. New York: Thomas Y. Crowell.
- Fodor, J. A. (1980). Methodological solipsism considered as a research strategy in cognitive psychology. *Behavioral and Brain Sciences*, 3, 63–109.
- Fodor, J. A. (1983). *The Modularity of Mind*. Cambridge, MA: MIT Press.
- Freeman, W. J. (2000). Emotion is Essential to All Intentional Behaviors. In M. D. Lewis & I. Granic (Eds.), *Emotion, Development, and Self-Organization: Dynamic Systems Approaches to Emotional Development* (pp. 209–235). Cambridge: Cambridge University Press.

- Friedman, B. H. (2010). Feelings and the body: The Jamesian perspective on autonomic specificity in emotion. *Biological Psychology*, 84, 383–393.
- Gagnon, K. T., Geuss, M. N., & Stefanucci, J. K. (2013). Fear influences perceived reaching to targets in audition, but not vision. *Evolution and Human Behavior*, 34, 49–54.
- Gagnon, S. & Doré, F. Y. (1994). Cross-sectional study of object permanence in domestic puppies (*Canis familiaris*). *Journal of Comparative Psychology*, 108(3), 220–232.
- Gendron, M. & Barrett, L. F. (2009). Reconstructing the Past: A Century of Ideas About Emotion in Psychology. *Emotion Review*, 1(4), 316–339.
- Gibson, E. J. & Walk, R. D. (1960). The "Visual Cliff". *Scientific American*, 202(4), 64–71.
- Gibson, J. J. (1966). *The Senses Considered as Perceptual Systems*. New York: Houghton Mifflin.
- Gibson, J. J. (1976). The Myth of Passive Perception: A Reply to Richards. *Philosophy and Phenomenological Research*, 37(2), 234–238.
- Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. New York: Psychology Press.
- Ginsburg, S. & Jablonka, E. (2007a). The Transition to Experiencing: I. Limited Learning and Limited Experiencing. *Biological Theory*, 2(3), 218–230.
- Ginsburg, S. & Jablonka, E. (2007b). The Transition to Experiencing: II. The Evolution of Associative Learning Based on Feelings. *Biological Theory*, 2(3), 231–243.
- Ginsburg, S. & Jablonka, E. (2010a). Experiencing: A Jamesian Approach. *Journal of Consciousness Studies*, 17(5-6), 102–124.

- Ginsburg, S. & Jablonka, E. (2010b). The evolution of associative learning: A factor in the Cambrian explosion. *Journal of Theoretical Biology*.
- Godfrey-Smith, P. (2016a). Individuality, subjectivity, and minimal cognition. *Biology & Philosophy*, 31(6), 775–795.
- Godfrey-Smith, P. (2016b). *Other Mind: The Octopus, the Sea, and the Deep Origins of Consciousness*. New York, NY: Farrar, Strauss and Giroux.
- Goldie, P. (1999). How We Think of Others' Emotions. *Mind & Language*, 14(4), 394–423.
- Goldie, P. (2000). *The Emotions: A Philosophical Exploration*. New York: Oxford University Press.
- Goldie, P. (2002). Emotions, feelings and intentionality. *Phenomenology and the Cognitive Sciences*, 1, 235–254.
- Goldie, P. (2004). Emotion, Feeling, and Knowledge of the World. In R. C. Solomon (Ed.), *Thinking about Feeling: Contemporary Philosophers on Emotions* (pp. 91–106). New York, NY: Oxford University Press.
- Goldie, P. (2009). Getting Feelings into Emotion Experience in the Right Way. *Emotion Review*, 1, 232–239.
- Goodale, M. A. (1995). The Cortical Organization of Visual Perception and Visuomotor Control. In S. M. Kosslyn & D. N. Osherson (Eds.), *An Invitation to Cognitive Science: Volume 2: Visual Cognition* (pp. 167–213). Cambridge, MA: MIT Press.
- Goodale, M. A. & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neuroscience*, 15, 20–25.
- Gordon, R. M. (1978). Emotion Labelling and Cognition. *Journal for the Theory of Social Behaviour*, 8(2), 125–135.

- Gottfried, G. M. & Jow, E. E. (2003). "I just talk with my heart": the mind-body problem, linguistic input, and the acquisition of folk psychological beliefs. *Cognitive Development*, 18, 79–90.
- Griffiths, P. E. (1990). Modularity, and the Psychoevolutionary Theory of Emotion. *Biology and Philosophy*, 5, 175–196.
- Griffiths, P. E. (1997). *What Emotions Really Are: The Problem of Psychological Categories*. Chicago: University of Chicago Press.
- Griffiths, P. E. (2004a). Emotions as Natural and Normative Kinds. *Philosophy of Science*, 71, 901–911.
- Griffiths, P. E. (2004b). Is Emotion a Natural Kind? In R. C. Solomon (Ed.), *Thinking about Feeling: Contemporary Philosophers on Emotions* (pp. 233–249). Oxford University Press.
- Griffiths, P. E. (2013). Current Emotion Research in Philosophy. *Emotion Review*, 5(2), 215–222.
- Hatfield, G. (2007). Did Descartes Have a Jamesian Theory of the Emotions? *Philosophical Psychology*, 20, 413–440.
- Hatfield, G. (2009a). Objectivity and Subjectivity Revisited: Color as a Psychobiological Property. In *Perception & Cognition: Essays in the Philosophy of Psychology* (pp. 281–296). Oxford, UK: Oxford University Press.
- Hatfield, G. (2009b). What Can the Mind Tell Us about the Brain? In *Perception and Cognition: Essays in the Philosophy of Psychology* (pp. 434–455). Oxford, UK: Oxford University Press.
- Hempel, C. G. (1980). The Logical Analysis of Psychology. In N. Block (Ed.), *Readings in*

- Philosophy of Psychology: Volume One* (pp. 14–23). Cambridge, MA: Harvard University Press. (Original work published 1949).
- Hendry, R. F. (2006). Elements, Compounds, and Other Chemical Kinds. *Philosophy of Science*, 73(5), 864–875.
- Hietanen, J. K., Glerean, E., Hari, R., & Nummenmaa, L. (2016). Bodily maps of emotions across child development. *Developmental Science*, 19(6), 1111–1118.
- Hills, P. J., Werno, M. A., & Lewis, M. B. (2011). Sad people are more accurate at face recognition than happy people. *Consciousness and Cognition*, 20(4), 1502–1517.
- Öhman, A. (2009). Of snakes and faces: An evolutionary perspective on the psychology of fear. *Scandinavian Journal of Psychology*, 50(6), 543–552.
- Holt, E. B. (1914). *The Concept of Consciousness*. New York, NY: Macmillan.
- Holt, E. B. (1931). *Animal Drive and the Learning Process: An Essay Toward Radical Empiricism*. New York, NY: Henry Holt.
- Ice Cube (1992). Check Yo Self. In *The Predator*. Glendale, CA: Lench Mob.
- Irons, D. (1894). Prof. James' Theory of Emotion. *Mind*, 3(9), 77–97.
- Izard, C. (1977). *Human Emotions*. New York: Plenum Press.
- Izard, C. (2007). Basic Emotions, Natural Kinds, Emotion Schemas, and a New Paradigm. *Perspectives on Psychological Science*, 2, 260–280.
- James, W. (1884). What is an Emotion? *Mind*, 9, 188–205.
- James, W. (1885). On the Function of Cognition. *Mind*, 10, 27–44.
- James, W. (1890). *The Principles of Psychology*, volume I & II. New York: Dover.

- James, W. (1894). The Physical Basis of Emotion. *Psychological Review*, 1, 516–529.
- James, W. (1992). Psychology: Briefer Course. In G. E. Myers (Ed.), *William James: Writings 1878-1899* (pp. 1–443). The Library of America. (Original work published 1892).
- Jeannerod, M. (2005). How Do We Decipher Others' Minds? In J.-M. Fellous & M. A. Arbib (Eds.), *Who Needs Emotions? The Brain Meets the Robot* (pp. 147–169). Oxford, UK: Oxford University Press.
- Johnson-Laird, P. N. & Oatley, K. (2008). Emotions, Music, and Literature. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *The Handbook of Emotions* (pp. 102–113). New York, NY: The Guildford Press, 3 edition.
- Kenny, A. (1963). *Action, emotion and will*. London, UK: Routledge & Kegan Paul.
- Kieran, M. (2010). Emotions, Art, and Immorality. In P. Goldie (Ed.), *The Oxford Handbook of Philosophy of Emotion* (pp. 681–703). Oxford, UK: Oxford University Press.
- Koffka, K. (1935). *Principles of Gestalt Psychology*. New York: Harbinger.
- Koriat, A. (2006). Are We Frightened Because We Run Away? Some Evidence from Metacognitive Feelings. In B. Uttl, N. Ohta, & A. L. Siegenthaler (Eds.), *Memory and Emotion: Interdisciplinary Perspectives* (pp. 83–103). Maiden, MA: Blackwell.
- Koriat, A., Ma'ayan, H., & Nussinson, R. (2006). The Intricate Relationships Between Monitoring and Control in Metacognition: Lessons for the Cause-and-Effect Relation Between Subjective Experience and Behavior. *Journal of Experimental Psychology: General*, 135(1), 36–69.
- Kornblith, H. (1993). *Inductive Inference and Its Natural Ground: An Essay in Naturalistic Epistemology*. Cambridge, MA: MIT Press.

- Laird, J. D. & Bresler, C. (1990). William James and the Mechanisms of Emotional Experience. *Personality and Social Psychology Bulletin*, 16, 636–651.
- Laird, J. D. & Bresler, C. (1992). The process of emotional experience: A self-perception theory. In M. S. Clark (Ed.), *Emotion: Volume 13, Review of personality and social psychology* (pp. 213–234). Newbury Park, CA: Sage.
- Landis, C. & Hunt, W. A. (1939). *The startle pattern*. New York: Farrar, Straus & Giroux.
- Lang, P. (1994). The Varieties of Emotional Experience: A Meditation on James-Lange Theory. *Psychological Review*, 101, 211–221.
- Lange, C. G. (1885). *Om Sindsbevaegelser*. Leipzig: Theodor Thomas.
- Lange, C. G. (1922). The Emotions. In C. G. Lange & W. James (Eds.), *The Emotions* (pp. 33–90). Williams & Wilkins Company.
- Langley, J. N. (1903). The Autonomic Nervous System. *Brain: A Journal of Neurology*, 26, 1–26.
- Larsen, J. T., Berntson, G. G., Poehlmann, K. M., Ito, T. A., & Cacioppo, J. T. (2008). The Psychophysiology of Emotion. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *The Handbook of Emotion* (pp. 180–195). New York: The Guildford Press, 3 edition.
- Laurence, S. & Margolis, E. (2002). Radical Concept Nativism. *Cognition*, 86, 25–55.
- Lazarus, R. S. (1982). Thoughts on the Relations Between Emotion and Cognition. *American Psychologist*, 37, 1019–1024.
- Lazarus, R. S. (1984). On the Primacy of Cognition. *American Psychologist*, 39, 124–129.
- Lazarus, R. S. (1991). *Emotion and Adaptation*. New York: Oxford University Press.

- Lazarus, R. S. (2001). Relational Meaning and Discrete Emotions. In K. R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal Processes in Emotion* (pp. 37–67). New York: Oxford University Press.
- Le, Q. V., Isbell, L. A., Matsumoto, J., Nguyen, M., Hori, E., Maior, R. S., Tomas, C., Tran, A. H., Ono, T., & Nishijo, H. (2013). Pulvinar neurons reveal neurobiological evidence of past selection for rapid detection of snakes. *Proceedings of the National Academy of Sciences*, 110(47), 19000–19005.
- Leder, H., Tinio, P. P. L., & Bar, M. (2011). Emotional valence modulates the preference for curved objects. *Perception*, 40, 649–655.
- LeDoux, J. E. (1996). *The Emotional Brain: The Mysterious Underpinnings of Emotional Life*. New York: Simon & Schuster.
- LeDoux, J. E. (2002a). Emotion, Memory and the Brain. *Scientific American*, 12, 62–71.
- LeDoux, J. E. (2002b). *Synaptic Self: How Our Brains Become Who We Are*. New York: Penguin Books.
- LeDoux, J. E. (2007). Unconscious and conscious contributions to the emotional and cognitive aspects of emotions: a comment on Scherer's view of what an emotion is. *Social Science Information*, 46(3), 395–405.
- LeDoux, J. E. (2008). Emotional colouration of consciousness: how feelings come about. In L. Weiskrantz & M. Davies (Eds.), *Frontiers of Consciousness: Chichele Lectures* (pp. 69–130). Oxford: Oxford University Press.
- LeDoux, J. E. (2014). Coming to terms with fear. *Proceedings of the National Academy of Science*, 111(8), 2871–2878.

- Leibovich, T., Cohen, N., & Henik, A. (2016). Itsy bitsy spider? Valence and self-relevance predict size estimation. *Biological Psychology*, 121, 138–145.
- Lewis, D. K. (1980). Mad Pain and Martian Pain. In N. Block (Ed.), *Readings in the Philosophy of Psychology* (pp. 216–222). Harvard University Press.
- Lyons, W. (1999). The philosophy of cognition and emotion. In T. Dagleish & M. Power (Eds.), *Handbook of Cognition and Emotion* (pp. 21–44). New York: John Wiley & Sons.
- MacLeod, M. (2013). Limitations of Natural Kind Talk in the Life Sciences: Homology and Other Cases. *Biological Theory*, 7, 109–120.
- Malebranche, N. (1674/1997). *The Search after Truth*. Cambridge: Cambridge University Press.
- Mandler, G. (1984). *Mind and Body: Psychology of Emotion and Stress*. New York: W. W. Norton & Company.
- Markus, H. (1990). On Splitting the Universe. *Psychological Science*, 1(3), 181–185.
- Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. Cambridge, MA: MIT Press.
- Martin, A. (2014). *How We Hope: A Moral Psychology*. Princeton, NJ: Princeton University Press.
- Miller, H. C., Gipson, C. D., Vaughan, A., Rayburn-Reeves, R., & Zentall, T. R. (2009). Object permanence in dogs: Invisible displacement in a rotation task. *Psychonomic Bulletin & Review*, 16(1), 150–155.
- Milner, A. D. & Goodale, M. A. (1995). *The visual brain in action*. Oxford: Oxford University Press.

- Morton, B. (2008). The Evolution of Eyes in the Bivalvia: New Insights. *American Malacological Bulletin*, 26(1-2), 35–45.
- Nagel, T. (1974). What Is It Like to Be a Bat? *The Philosophical Review*, 83(4), 435–450.
- Neu, J. (2010). An Ethics of Emotion? In P. Goldie (Ed.), *The Oxford Handbook of Philosophy of Emotion* (pp. 501–517). Oxford, UK: Oxford University Press.
- Niedenthal, P. (2008). Emotion Concepts. In M. Lewis, J. Haviland-Jones, & L. Barrett (Eds.), *Handbook of Emotions*, volume 3 (pp. 587–600). New York: Guildford Press, 3 edition.
- Niedenthal, P., Wood, A., & Rychlowska, M. (2014). Embodied Emotion Concepts. In L. Shapiro (Ed.), *The Routledge Handbook of Embodied Cognition* (pp. 240–249). Abingdon, UK: Routledge.
- Nummenmaa, L., Glerean, E., Hari, R., & Hietanen, J. K. (2014). Bodily maps of emotions. *Proceedings of the National Academy of Sciences*, 111(2), 646–651.
- Nussbaum, M. C. (2001). *Upheavels of Thought: The Intelligence of the Emotions*. Cambridge: Cambridge University Press.
- Nussbaum, M. C. (2004). *Hiding from Humanity: Disgust, Shame, and the Law*. Princeton, NJ: Princeton University Press.
- Ortony, A., Clore, G. L., & Collins, A. (1988). *The Cognitive Structure of Emotions*. Cambridge, UK: Cambridge University Press.
- Panichello, M. F., Kveraga, K., Chaomon, M., Bar, M., & Barrett, L. F. (2017). Internal valence modulates the speed of object recognition. *Scientific Reports*, 7(1), 1–10.
- Panksepp, J. (1998). *Affective neuroscience: the foundations of human and animal emotions*. New York, NY: Oxford University Press.

- Perry, J. (2017). Identity and Self-Knowledge. *Philosophy, Theory, and Practice in Biology*, 9(5), 1–12.
- Plutchik, R. (1970). Emotions, Evolution, and Adaptive Processes. In M. B. Arnold (Ed.), *Feelings and Emotions: The Loyola Symposium* (pp. 3–24). Academic Press.
- Prinz, J. J. (2002). *Furnishing the Mind: Concepts and Their Perceptual Basis*. Cambridge, MA: MIT Press.
- Prinz, J. J. (2003). Emotion, Psychosemantics, and Embodied Appraisals. In A. Hatzimoysis (Ed.), *Philosophy and the Emotions*, volume 52 of *Royal Institute of Philosophy Supplement* (pp. 69–86). Cambridge: Cambridge University Press.
- Prinz, J. J. (2004). *Gut Reactions: A Perceptual Theory of Emotion*. Oxford: Oxford University Press.
- Prinz, J. J. (2012). *The Conscious Brain: How Attention Engenders Experience*. New York, NY: Oxford University Press.
- Rage Against The Machine (1992). Freedom. In *Rage Against The Machine*. Los Angeles, CA: Epic Records.
- Ratcliffe, M. (2005). William James on Emotion and Intentionality. *International Journal of Philosophical Studies*, 13(2), 179–202.
- Reisenzein, R. (1983). The Schachter Theory of Emotion: Two Decades Later. *Psychological Bulletin*, 94(2), 239–264.
- Reisenzein, R. (2017). Varieties of Cognition-Arousal Theory. *Emotion Review*, 9(1), 17–26.
- Reisenzein, R. & Döring, S. A. (2009). Ten Perspectives on Emotional Experience: Introduction to the Special Issue. *Emotion Review*, 1(3), 195–205.

- Reisenzein, R., Meyer, W.-U., & Schützwohl, A. (1995). James and the Physical Basis of Emotion: A Comment on Ellsworth. *Psychological Review*, 102, 757–761.
- Robinson, J. (1995). Startle. *The Journal of Philosophy*, 92, 53–74.
- Robinson, J. (2005). *Deeper Than Reason: Emotion and its Role in Literature, Music, and Art*. New York: Oxford University Press.
- Robinson, J. (2010). Emotional Responses to Music: What Are They? How Do They Work? And Are They Relevant to Aesthetic Appreciation? In P. Goldie (Ed.), *The Oxford Handbook of Philosophy of Emotion* (pp. 651–680). Oxford: Oxford University Press.
- Rorty, A. (1980). Explaining Emotions. In A. Rorty (Ed.), *Explaining Emotions* (pp. 103–126). Berkeley, CA: University of California Press.
- Rorty, A. (2004). Enough Already with "Theories of the Emotions". In R. C. Solomon (Ed.), *Thinking about Feeling: Contemporary Philosophers on Emotions* (pp. 269–278). New York, NY: Oxford University Press.
- Russell, J. A. (1980). A Circumplex Model of Affect. *Journal of Personality and Social Psychology*, 39(6), 1161–1178.
- Russell, J. A. (2003). Core Affect and the Psychological Construction of Emotion. *Psychological Review*, 110(1), 145–172.
- Russell, J. A. (2017). Mixed Emotions Viewed from the Psychological Constructionist Perspective. *Emotion Review*, 9(2), 111–117.
- Salazar-Lopez, E., Dominguez, E., Juarez Ramos, V., de la Fuente, J., Meins, A., Iborra, O., Galvez, G., Rodriguez-Artacho, M. A., & Gomez-Milan, E. (2015). The mental and subjective skin: Emotion, empathy, feelings and thermography. *Consciousness and Cognition*, 34, 149–162.

- Scarantino, A. (2003). Affordances Explained. *Philosophy of Science*, 70(5), 949–961.
- Scarantino, A. (2009). Core Affect and Natural Affective Kinds. *Philosophy of Science*, 76(5), 940–957.
- Scarantino, A. (2010). Insights and Blindspots of the Cognitivist Theory of Emotions. *British Journal for the Philosophy of Science*, 61, 729–768.
- Scarantino, A. (2014). Basic Emotions, Psychological Construction, and the Problem of Variability. In L. F. Barrett & J. A. Russell (Eds.), *The Psychological Construction of Emotion* (pp. 334–376). New York, NY: The Guildford Press.
- Scarantino, A. & Griffiths, P. (2011). Don't Give Up on Basic Emotions. *Emotion Review*, 3(4), 444–454.
- Schachter, S. & Singer, J. E. (1962). Cognitive, Social, and Physiological Determinants of Emotional State. *Psychological Review*, 69(5), 379–399.
- Scherer, K. (1993). Studying the emotion-antecedent appraisal process: An expert system approach. *Cognition & Emotion*, 7, 325–355.
- Scherer, K. (1996). Emotion. In M. Hewstone, W. Stroebe, & G. M. Stephenson (Eds.), *Introduction to Social Psychology* (pp. 279–315). Oxford: Blackwell, 2 edition.
- Scherer, K. (1999). Appraisal Theory. In T. Dagleish & M. Power (Eds.), *Handbook of Cognition and Emotion* (pp. 637–664). New York: John Wiley & Sons.
- Scherer, K. (2001). The Nature and Study of Appraisal: A Review of the Issues. In K. R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal Processes in Emotion* (pp. 369–391). New York: Oxford University Press.

- Schorr, A. (2001). Appraisal: The Evolution of an Idea. In K. R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal Processes in Emotion* (pp. 20–34). New York, NY: Oxford University Press.
- Shibasaki, M. & Kawai, N. (2009). Rapid detection of snakes by Japanese monkeys (*Macaca fuscata*): An evolutionarily predisposed visual system. *Journal of Comparative Psychology*, 123(2), 131–135.
- Skinner, B. F. (1953). *Science and Human Behavior*. New York, NY: Macmillan Publishing Co.
- Slaby, J. (2008). Affective intentionality and the feeling body. *Phenomenology and the Cognitive Sciences*, 7, 429–444.
- Solomon, R. C. (1976). *The Passions: Emotions and the Meaning of Life*. Indianapolis: Hackett.
- Solomon, R. C. (1980). Emotions and Choice. In A. O. Rorty (Ed.), *Explaining Emotions* (pp. 251–281). Berkeley: University of California Press.
- Solomon, R. C. (2003). Emotions, Thoughts and Feelings: What is a 'Cognitive Theory' of the Emotions and Does it Neglect Affectivity? In A. Hatzimoysis (Ed.), *Philosophy and the Emotions*, volume 52 of *Royal Institute of Philosophy Supplement* (pp. 1–18). Cambridge: Cambridge University Press.
- Solomon, R. C. (2004). Emotions, Thoughts, and Feelings: Emotions as Engagements with the World. In R. C. Solomon (Ed.), *Thinking about Feeling: Contemporary Philosophers on Emotions* (pp. 76–88). New York, NY: Oxford University Press.
- Spencer, H. (1855). *The Principles of Psychology*. London: Longman, Brown, Green, and Longmans.

- Spencer, H. (1863). *Essays: scientific, political, and speculative*, volume 2. London, UK: Williams and Norgate.
- Strout, S. L., Sokol, R. I., Laird, J. D., & Thompson, N. S. (2004). The Evolutionary Foundation of Perceiving One's Own Emotions. *Behavior and Philosophy*, 32(2), 493–502.
- Susskind, J. M., Lee, D. H., Cusi, A., Feiman, R., Grabski, W., & Anderson, A. K. (2008). Expressing fear enhances sensory acquisition. *Nature Neuroscience*, 11(7), 843–850.
- Taylor, J. (2013). Emotional Sensations and the Moral Imagination in Malebranche. In H. M. Lloyd (Ed.), *The Discourse of Sensibility: The Knowing Body in the Enlightenment*, volume 35 of *Studies in History and Philosophy of Science* (pp. 63–84). Cham, Switzerland: Springer.
- Terfry, R. (2015). *Wicked and Weird: The Amazing Tales of Buck 65*. USA: Doubleday Canada.
- Titchener, E. B. (1909). *Lectures on the Experimental Psychology of Thought-Processes*. New York, NY: Macmillan.
- Todd, R. M., Talmi, D., Schmitz, T. W., Susskind, J., & Anderson, A. K. (2012). Psychophysical and Neural Evidence for Emotion-Enhanced Perceptual Vividness. *The Journal of Neuroscience*, 32(33), 11201–11212.
- Tollman, E. C. (1932). *Purposive Behavior in Animals and Men*. New York, NY: Century.
- Tomkins, S. (1962). *Affect, Imagery, and Consciousness: Volume 1: The positive affects*. New York: Springer.
- Tomkins, S. (1963). *Affect, Imagery, and Consciousness: Volume 2: The negative affects*. New York: Springer.
- Tooby, J. & Cosmides, L. (2008). The Evolutionary Psychology of the Emotions and Their Relationship to Internal Regulatory Variables. In M. Lewis, J. M. Haviland-Jones, & L. F.

- Barrett (Eds.), *Handbook of Emotions* (pp. 114–137). New York, NY: Guildford Press, 3 edition.
- Trestman, M. (2013). The Cambrian Explosion and the Origins of Embodied Cognition. *Biological Theory*, 8(1), 80–92.
- Triana, E. & Pasnak, R. (1981). Object permanence in cats and dogs. *Animal Learning & Behavior*, 9(1), 135–139.
- von Uexküll, J. (1926). *Theoretical Biology*. New York, NY: Barcourt, Brace & Co.
- Watson, J. B. (1930). *Behaviorism*. New York: W. W. Norton & Company.
- Zadra, J. R. & Clore, G. L. (2011). Emotion and Perception: The Role of Affective Information. *Wiley Interdisciplinary Reviews: Cognitive Science*, 2, 675–685.
- Zagalo, N., Torres, A., & Branco, V. (2005). Emotion Spectrum Developed by Virtual Storytelling. In G. Subsol (Ed.), *Virtual Storytelling: Using Virtual Reality Technologies for Storytelling*, volume 3805 of *Lecture Notes in Computer Science*. Berlin: Springer.
- Zajonc, R. B. (1980). Feeling and Thinking: Preferences Need No Inferences. *American Psychologist*, 35, 151–175.
- Zajonc, R. B., Murphy, S. T., & McIntosh, D. N. (1993). Brain Temperature and Subjective Emotional Experience. In M. Lewis & J. Haviland (Eds.), *Handbook of Emotions* (pp. 209–220). New York: Guildford Press, 1 edition.